THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Critical Fluctuations and Coupling of Stochastic Neural Mass Models

Matthew J. Aburn

BSc, BA, MInfTech

A thesis submitted for the degree of Doctor of Philosophy at

The University of Queensland in 2016

School of Mathematics and Physics

Abstract

Neuroimaging and implanted medical devices give a view onto large scale oscillations in the brain that rapidly change in both power and synchrony at different frequencies. These measured time series reveal a projection to just one dimension for each location, a moving shadow of the complex underlying activity. For this reason mathematical models of the activity are essential to make full use of the time series. The flexible onset and cessation of oscillations may have an important role in allowing the brain rapidly to rearrange its effective network structure. The onset of pathological oscillations is also central to neurological disorders such as epilepsy and Parkinson's disease. It is in the regime close to linear instability that transitions to and from oscillation can readily occur, by mechanisms of bifurcation and multistability. Empirical evidence of *critical fluctuations* and *critical slowing down* in neural dynamics also suggest proximity to linear instability.

This thesis develops novel methods to analyze stochastic dynamical models near the point of transition where oscillations start and stop, allowing to relate transitions in the models to changes in time series statistics that are accessible in clinical and experimental recordings.

Taking as an example a widely used biophysical model of collective neural dynamics (the Jansen-Rit model) we show that near the point of oscillation onset at a supercritical Hopf bifurcation there can be a dramatic increase in autocorrelation length as well as power law scaling over four orders of magnitude in fluctuation size and duration, but that these time series indicators will depend sensitively on the direction in state space of input fluctuations and hence on which neuronal subpopulation is stochastically perturbed.

In general near a supercritical Hopf bifurcation non-circular oscillations emerge on an arbitrarily curved two-dimensional surface embedded within the N-dimensional state space. This makes it difficult to determine exactly how noise will affect oscillation phase and amplitude. By applying Poincaré normal form transformations to find a parameter-dependent coordinate system in which the unperturbed oscillations are simple circles in a flat plane, then using Stratonovich stochastic calculus to transform noise perturbations to the same coordinate system, this makes explicit the effect of noise on phase and amplitude. We give a criterion for this reduced model to be a good approximation, balancing noise intensity and center manifold stability.

Optionally averaging the diffusion process in Fokker-Planck form gives a still simpler weak approximation of the oscillations in a symmetrical standard form that retains the leading order stochastic effects on phase and amplitude dynamics. We demonstrate by large scale numerical simulation of the stochastic differential equations that the oscillation time series statistics of interest are preserved by transformation to this symmetrical weak approximation.

Close to instability, dynamic synchrony of networked oscillations can depend on both phase and amplitude changes of constituent oscillators. In computational neuroscience studies of network synchrony near instability, coupled systems of Hopf bifurcation normal forms and Generalized Hopf normal forms have been used successfully as minimal models with local oscillation onset occurring via bifurcation and bistability respectively. Such abstract models are often parameterized somewhat arbitrarily, as the individual flow in the state space is qualitatively the same for a range of parameters. But this assumption breaks down when the systems are not isolated.

In this deterministic setting we demonstrate a second application of normal form transformations to constrain the parameters and coupling of a normal form network model based on a more detailed biophysical model. We show that using the Poincaré normal form allows the simple model to match the collective synchronization behavior of the more detailed biophysical model, whereas using the topological normal form with symmetrical diffusive coupling does not.

Ultimately the methods developed in this thesis allow to predict from a given biophysical model how time series statistics of oscillations will change as biological parameters are changed in experiment. This can assist in testing models empirically, to arrive at better models. After suitable models are evolved this will then allow diagnostic inference of brain state that cannot be observed directly, using clinical neuroimaging time series.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

Publications during candidature

Aburn, M. J., Holmes, C. A., Roberts, J. A., Boonstra, T. W., Breakspear, M. (2012). Critical fluctuations in cortical models near instability. *Frontiers in Physiology* **3**: 331.

Publications included in this thesis

Aburn, M. J., Holmes, C. A., Roberts, J. A., Boonstra, T. W., Breakspear, M. (2012). Critical fluctuations in cortical models near instability. *Frontiers in Physiology* **3**: 331 – incorporated as Chapter 2:

Contributor	Statement of contribution
Aburn, M. J. (Candidate)	Conceptual design (10%) Designed and implemented simulations (100%) Bifurcation and time series analyses (60%) Wrote and edited paper (50%)
Holmes, C. A.	Conceptual design (15%) Bifurcation and time series analyses (10%) Wrote and edited paper (10%)
Roberts, J. A.	Bifurcation and time series analyses (20%) Wrote and edited paper (10%)
Boonstra, T. W.	Conceptual design (15%) Wrote and edited paper (5%)
Breakspear, M.	Conceptual design (60%) Bifurcation and time series analyses (10%) Wrote and edited paper (25%)

Contributions by others to the thesis

Sections 2.3.5 and 2.4.2 (analysis of fluctuation scaling) are primarily the work of Dr James A. Roberts, who also contributed to critical review of drafts.

Dr Cathy A. Holmes conceived the initial idea of applying deterministic normal form transformations to estimate leading order stochastic terms.

Prof. Michael Breakspear contributed conceptual design of the project, critical review of drafts, jointly wrote Section 2.1 and partly authored code applied in Section 2.3.5 (analysis of fluctuation scaling).

Prof. dr. Andreas Daffertshofer conceived the idea of applying averaging in Fokker-Planck form.

Dr Tjeerd Boonstra contributed to conceptual design of the project and to review of drafts of Chapter 2.

Figures 1.2 and 1.3 are by Prof. Yuri A. Kuznetsov, reproduced under Creative Commons Attribution-NonCommercial-ShareAlike license.

Figures 3.1A and 3.1B are by D. G. Aronson, G. B. Ermentrout and N. Kopell. Reprinted from Aronson et al. (1990), with permission from Elsevier.

Statement of parts of the thesis submitted to qualify for the award of another degree

None.

Acknowledgments

I gratefully acknowledge my principal advisor Cathy Holmes, whose expertise and patience in discussion guided me to understand one thing after another; my associate advisor Michael Breakspear for his clear expositions of the most complex domain and generosity in handing to me many keys to success as a research scientist, though perhaps I did not grasp them; and Andreas Daffertshofer, whose instruction and encouragement were crucial in turning this attempt to understand one system into a general method more useful to others.

For help in discussing and checking this work and for suggesting valuable improvements many thanks are due to James Roberts, Kevin Burrage, Pamela Burrage, Thomas Taimre, John Terry, Phil Isaac, Tjeerd Boonstra and Steve Mehrkanoon. For the limitations that remain I am responsible. My deepest gratitude is to my wife Jaime, who supported and encouraged me throughout and sacrificed so that I could leave a career and take on this project.

Keywords

stochastic dynamical system, neural mass model, Hopf bifurcation, neuroimaging, synchronization

Australian and New Zealand Standard Research Classifications

ANZSRC code: 010204, Dynamical Systems in Applications, 50%

ANZSRC code: 110999, Neurosciences, 30%

ANZSRC code: 010406, Stochastic Analysis and Modelling, 20%

Fields of Research (FoR) Classification

FoR code: 0102, Applied Mathematics, 50%

FoR code: 1109, Neurosciences, 30%

FoR code: 0104, Statistics, 20%

Mathematical notation

Scalar and vector quantities

Scalars are in italic type: $x, f(x)$. Vectors and vector valued functions are in bold: $\mathbf{x},\ \mathbf{f}(\mathbf{x})$.

Vectors in $\mathbb{R}^n$ are written with square brackets,

$$\mathbf{x} = \begin{bmatrix} 1.5 \\ 2 \\ 3.83 \end{bmatrix}, \qquad \text{or written as the transpose of a row vector:} \quad (1.5,\ 2,\ 3.83)^{\mathrm{T}}.$$

Differential operators

A dot denotes differentiation with respect to time: $\dot{x} = \frac{dx}{dt}$.

$D\mathbf{f}$ denotes the Jacobian matrix differential of a function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$,

so $A = D\mathbf{f}(\mathbf{x})$ is a matrix with entries $A_{ij} = \frac{\partial f_i}{\partial x_j}$ for $i = 1 \cdots m,\ j = 1 \cdots n$.

For a function of two vectors, $\mathbf{h} :\ \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$,

$\dfrac{\partial}{\partial \mathbf{w}} \mathbf{h}(\mathbf{v}, \mathbf{w})$ denotes partial differentiation with respect to one vector argument,

that is, it gives a matrix B with elements $B_{ij} = \frac{\partial}{\partial w_j} h_i(\mathbf{v}, \mathbf{w})$ for $i = 1 \cdots m,\ j = 1 \cdots n$.

The partial differential equation

$$\frac{\partial}{\partial t}\, p(x,t) = -\frac{\partial}{\partial x}(f(x,t)p(x,t)) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(g(x,t)^2 p(x,t)),$$

will be written using an algebra of differential operators:

$$\partial_t\, p(x,t) = \left[-\partial_x f(x,t) + \frac{1}{2}\partial_x^2 g(x,t)^2\right] p(x,t),$$

where multiplication takes precedence over differentiation.

Stochastic differential equations (SDE)

Ito equations are written: $\quad dx = f(x,t)\, dt + g(x,t)\, dW.$

Stratonovich equations are written with a circle: $\quad dx = f(x,t)\, dt + g(x,t) \circ dW.$

For a d-dimensional multivariate SDE $\mathbf{dx} = \mathbf{f}(\mathbf{x}, t)\, dt + G(\mathbf{x}, t)\, \mathbf{dW}$, G is a $d \times p$ matrix of coefficients for p independent Wiener processes informally denoted $\mathbf{dW} = (dW_1, dW_2, \cdots, dW_p)$.

Contents

Chapter 1

Introduction

Changes in neural systems near the onset of oscillations are the focus of this work, because of the importance of these transitions to brain function. In this introduction we first review the biophysical basis for oscillations in the complex fluctuating system of the brain, then introduce some of the mathematical tools used to distill these in minimal models and to understand them.

1.1 The dynamic brain

To survive, the human brain requires flexibility to change, both to maintain internal representations of the outside world, and to respond quickly to changes in input, ultimately including rapid changes of action by transitions in the motor control system. But the brain must also persist. Therefore the dynamical mechanisms of brain function are neither strongly stable nor unstable but remain in the regime close to instability where transitions occur easily. It has been speculated that this necessary balance in the brain is realized by mechanisms of self-organized criticality (Linkenkaer-Hansen et al., 2001; Chialvo, 2010; Rubinov et al., 2011; Shew and Plenz, 2013) though this remains to be settled (Bedard et al., 2006; Touboul and Destexhe, 2010).

These transitions in macroscopic state can potentially occur in several ways:

- High-dimensional systems can undergo discontinuous or continuous *phase transitions* in which the qualitative change is across multiple space and time scales (Binney, 1992).

- By contrast a *bifurcation* is a transition in the low-dimensional dynamics of macroscopic variables between qualitatively different behaviors as a parameter is changed continuously. Bifurcation theory begins with the repertoire of transitions for autonomous systems without noise (Guckenheimer and Holmes, 1985; Kuznetsov, 2010).
- The phenomenon of *intermittency* shows that even in this autonomous, noise-free setting spontaneous transitions between behaviors can occur without a parameter being changed, for example in a chaotic system with near-periodic orbits (Pomeau and Manneville, 1980; Goodfellow and Glendinning, 2013).
- Often a system will simultaneously have more than one macroscopic stable state or attractor. This *bistability* or *multistability* means that as soon as fluctuations are added to the picture, the system may switch between these attractors, and will switch more frequently as the stability of an attractor is weakened.

In any high dimensional dissipative system, fluctuations occur. These are represented in macroscopic models by noise processes. In such stochastic dynamical systems, near the point of bifurcation phenomena arise such as *critical slowing down* and *critical fluctuations* that resemble those near high-dimensional phase transitions (Scheffer et al., 2009; Kelso, 2010; Kuehn, 2011). Haken made this analogy concrete by connecting bifurcations of stochastic macroscopic dynamics with phase transitions in systems far from thermodynamic equilibrium, in terms of Ginzburg-Landau theory (Graham and Haken, 1970; Haken, 1978).

Oscillations are thought to play a key role in the brain's flexible macroscopic transitions. At microscopic scales, individual neurons fire stochastically with discrete spike trains. But at mesoscopic (10^{-4}m to 10^{-3}m) and macroscopic ($> 10^{-3}$m) scales dynamically changing oscillations at frequencies from 0.1Hz to 200Hz are evident in the collective electrophysiological activity (Buzsáki, 2006). Clear transitions in synchrony are observed between oscillations in widely separated brain regions (Engel et al., 2001; Boonstra et al., 2009b) as well as cross-frequency coupling (Jirsa and Müller, 2013). Since phase synchronization can gate communication between regions (Womelsdorf et al., 2007; Bastos et al., 2015), dynamic changes in oscillation power and synchrony locally and between brain regions allow a rapid dynamic rearrangement of the effective network structure of the brain, even

as the structural network of neuron-to-neuron connections stays fixed.[1]

The flip side of this physiological role of flexible transitions of oscillation is that in neurological disorders including epilepsy and Parkinson's disease the onset of pathological oscillations can be a defining part of the disease. For example transitions to pathological oscillations are characteristic of epileptic seizure (Wendling et al., 2016) and changes in synchrony may facilitate ictal propagation between cortical regions as one aspect of the epileptogenic process (Mormann et al., 2005; Schmidt et al., 2016).

These are the reasons why this thesis engages oscillations in mathematical models of neural systems, and why there is a particular focus on the regime close to linear instability where bifurcations allow the onset and cessation of oscillations.

In the rest of this chapter we introduce in greater depth neural oscillations, noise and bifurcations and how computational neuroscience elucidates these aspects of neural dynamics.

1.2 Oscillations in the nervous system

Oscillations, or periodically repeating waveforms, are prominent features of electrophysiological activity in the nervous system at mesoscopic and macroscopic spatial scales, with the frequencies thought to be important in brain function ranging from 0.1 Hz to 200 Hz (Buzsáki and Draguhn, 2004).

Non-invasive neuroimaging

The large-scale oscillations were recognized as early as Berger (1929), when the first non-invasive measurements via EEG (electroencephalography) were measured at the scalp. EEG measures the difference in electric potential between each measurement electrode and a reference electrode (often also on the scalp). In modern experimental and clinical settings a conducting gel and slight pressure are enough to ensure good electrical contact with skin tissue. The field being measured here is primarily due to currents induced by post-synaptic potentials across cellular membranes at the apical dendrites of pyramidal neurons, particularly in layers 3-5 of the folded cortical sheet, as these large dendrites have a predominant orientation perpendicular to the sheet.

[1] This structural network also dynamically reorganizes through synaptic and dendritic plasticity and through axonal rewiring, but on longer time scales ($> 10^0$ s).

Just one or two measurement electrodes and a reference electrode are sufficient to clearly observe oscillations in the traditionally named delta (0.5-3 Hz), theta (4-7 Hz) and alpha (8-13 Hz) frequency ranges by examining a plot of the raw signals by eye. This has immediate clinical application in scoring sleep quality by distinguishing the different stages of sleep. Higher frequency bands are traditionally labeled beta (13-30 Hz) and gamma (> 30 Hz). For neuroimaging in clinical and experimental settings a full EEG cap with between 19 and 64 electrodes at standardized locations is used (Nuwer et al., 1998).

Alpha oscillations are the most readily observable. These are strong especially in occipital regions when the subject is in an eyes-closed resting state, and modulate with visual attention (Niedermeyer and Silva, 2005). Beta oscillations are strongest in the somatosensory and primary motor cortex areas, and to a lesser extent occipital areas (Jensen et al., 2005).

Neuroscience and clinical neurology have an extensive body of knowledge that uses the above traditional naming of frequency bands. But this does not mean that these bands are optimal for neuroscience signal analysis in all cases. Shackman et al. (2010) proposed a method to divide the spectrum of oscillations into bands with maximally independent information, applied for example in Chowdhury et al. (2014).

Scalp EEG has poor spatial resolution (> 3 cm) as the fields measured arise from the summation of the activity of neurons over a wide area, as spatial mixing results from volume conduction of ions in the brain tissue and the interaction of the electric field with the skull (Nunez and Srinivasan, 2006). This can be partially improved with source reconstruction using lead field models (Pascual-Marqui et al., 1994). So the sources of oscillations recorded with EEG cannot be localized precisely in the brain. On the other hand EEG has excellent temporal resolution with sample rate up to 2000 Hz.

For high time resolution neuroimaging there is also MEG (magnetoencephalography) which has similar temporal resolution to EEG but provides higher spatial resolution.

EMG (electromyography) measures the changing electric field due to muscle activity, and because unlike in the brain there is here a one-to-one relationship between an action potential of an individual alpha motor neuron in the spine and the contraction of an motor unit substructure within a muscle, we can infer from EMG the activity of a subset of the spinal neural populations.

While fMRI (functional magnetic resonance imaging) provides a high spatial resolution ($\sim$ 2 mm) view of neural oxygen metabolism, but its time resolution is fundamentally limited by physiology to time scales $>$ 1 s so it is not able to measure the oscillatory behavior described here. Recently studies recording EEG and fMRI simultaneously allow combining the strengths of both modalities.

Invasive recordings

Invasive recordings either during surgery or from electrodes on implanted medical devices allow much finer localization of the signal being measured, while retaining temporal resolution.

Sources of these recordings of the intercellular local field potential include electrocorticographic grids (ECoG) implanted inside the cranium on the surface of the cortex, either inside or outside the dura mater. These can provide recording from a grid of 64 to 252 electrodes distributed over several centimeters of cortical surface (Rubehn et al., 2009).

Stereotactic EEG (SEEG, or depth-EEG) records using multiple 0.8 mm electrodes inserted deep in cortical and subcortical neural tissues, with each electrode having 5-18 separate measurement contacts along its length (Cosandier-Rimele et al., 2007). Using multi-modal imaging and examination of the signal at each channel the precise anatomical location of each contact within the brain can be determined.

Recording during surgery using multiple microelectrodes is common for example for patients being treated with deep brain stimulation (DBS) (Kinfe and Vesper, 2013), and chronically implanted DBS devices can also include recording microelectrodes.

In in vivo animal experiments chronically implanted microscopic tetrode arrays allow measurement of the intercellular electric field down to the resolution of several neurons, where individual neuronal activity can be disambiguated statistically (Ecker et al., 2010).

What is recorded by these methods as “local field potential” can vary depending on the spatial scale of the recording, which in turn is partly determined by electrode size (Nunez et al., 2001).

1.2.1 Neuroimaging time series and analysis methods

In both clinical and experimental application, a series of signal processing and analysis techniques are commonly applied to EEG and MEG neuroimaging time series data. These methods, and the basic nature of each EEG/MEG channel as a 1-dimensional measurement function of an n-dimensional internal state, inform the choices of 1-dimensional projection and signal processing that we shall apply to simulated time series in Chapters 2 and 5 of this thesis, so that findings about time series statistics of stochastic oscillations remain comparable to time series typically obtainable in neuroimaging.

Preprocessing

A (digital) narrowband notch filter is usually applied to remove artifacts from power line interference (either 50 or 60 Hz) and its harmonics at integer multiples of that frequency. Slowly changing resistance of an electrode contact with the skin, for example due to sweating, can introduce slow time drift artifacts in EEG time series. Despite this, applying a high-pass filter or subtracting a least-squares linear trend is best deferred to later analysis steps, as low frequency components can also be physiologically meaningful. The instantaneous mean across channels is often subtracted from all channels to re-reference to an average baseline (Pascual-Marqui and Lehamann, 1993). Noisy channels may be detected and omitted in computation of that average. This re-referencing is particularly important if assessing oscillation synchrony as it avoids spurious correlations between channels due to the common reference signal (Fein et al., 1988). The electrophysiology of muscle activity causes obvious artifacts in EEG data, particularly eye blinks. Either manual scoring or automated signal processing techniques such as ICA (independent component analysis) are usually used to remove these artifacts (Delorme and Makeig, 2004).

Time resolved frequency and power

For each channel, the signal $x(t)$ may be restricted to examine a small interval of time by multiplying it by a *window function* centered on a time τ. Similarly, the signal may be restricted to examine a narrow frequency range, by applying a band-pass filter centered on a frequency f. But in neuroimaging we are concerned with oscillations whose power and phase relationships change dynamically in time, at different frequencies. Therefore it is useful to have a complex-valued function $s(f, t)$ for each channel that estimates the amplitude and phase of different frequency components f and also shows how these change with time t.

However any definition of a function $s(f,t)$ that gives a such view of the signal resolved in both frequency and time must involve a tradeoff between time resolution and frequency resolution, as the product of these is fundamentally bounded by an uncertainty principle (Gröchenig, 2001). In practice the time-resolved complex spectrum $s(f,t)$ is estimated via either a short-time Fourier transform, which uses a common moving window in time for all frequencies, or by continuous wavelet transforms, which typically use a higher time resolution for higher frequencies.

In both cases the time-resolved power spectral density (PSD) $|s(f,t)|^2$ gives the distribution of signal power among frequencies. In cases where the signal is approximately stationary, the PSD across the whole time series is estimated by the method of Welch (1967), averaging over overlapping time windows.

The Hilbert transform $\mathcal{H}$ is used to extend the real-valued signal $x(t)$ of each channel to a complex-valued function of time called the *analytic signal*: $a(t) = x(t) + \mathrm{i}\,\mathcal{H}(x)(t)$.
Here $\mathcal{H}$ is most simply defined as multiplication by $-\mathrm{i}\,\mathrm{sgn}(f)$ in the frequency domain:

$$\widehat{\mathcal{H}(x)}(f) = -\mathrm{i}\,\mathrm{sgn}(f)\,\hat{x}(f), \tag{1.1}$$

where $\hat{x}$ denotes the Fourier transform of x and $\mathrm{sgn}(f)$ is the signum function. Since $1 + \mathrm{i}(-\mathrm{i}\,\mathrm{sgn}(f)) = 2\theta(f)$, where θ is the Heaviside step function, it follows from the linearity of the Fourier transform that $a(t)$ is related to $x(t)$ by simply setting the negative frequency components of $x(t)$ to zero (and doubling the result) (Bruns, 2004). From this analytic signal $a(t)$ we can then define the instantaneous amplitude $|a(t)|$, instantaneous phase $\arg(a(t))$ and instantaneous power $|a(t)|^2$ of the signal (Tass et al., 1998). The terms "Hilbert amplitude" and "Hilbert phase" are used with these meanings.

In neuroimaging, band-pass filters are sometimes applied first to separate different oscillations, then the analytic signal for each frequency band calculated, yielding its instantaneous amplitude and phase. The close relationship between short-time Fourier transform, complex wavelet transforms, and the Hilbert transform of band-passed signals is proved in Bruns (2004).

The above use of the Hilbert transform in neuroimaging is one reason it will be applied in the computational studies of Chapters 2 and 6, where we choose to define the oscillation amplitude of simulated

models by the instantaneous amplitude of a 1-dimensional projection of the state, thus making it comparable to oscillation amplitude measured in experimental time series.

Relationships between channels

Rather than time-frequency activity in individual brain regions, what is of most interest in neuroimaging is interaction between regions in the network of the brain, including phase relationships of oscillations within and between frequency bands. Here the methods employed in neuroscience for determining interactions from measured time series are diverse. The most common approach is to estimate statistical relationships between the data of separate channels.

To assess synchrony and phase coupling of oscillations, relationships are sometimes assumed to be unchanging over a particular time interval in an experiment (i.e. channels are assumed to be pairwise jointly stationary in the wide sense). *Coherency* is then the normalized cross-spectrum,

$$C_{ij}(f) = \frac{\langle \hat{x}_i(f) \hat{x}_j^*(f) \rangle}{\sqrt{\langle |\hat{x}_i(f)|^2 \rangle \langle |\hat{x}_j(f)|^2 \rangle}}, \tag{1.2}$$

a complex value for each pair of channels i, j at each frequency f. Its modulus (*coherence*) measures the linear correlation at each frequency between channels, ranging between 0 for linear independence and 1 for perfect correlation. For channels with significant coherence at frequency f, the phase of $C_{ij}(f)$ indicates the phase lag between channels at that frequency.

Coherence depends not only on phase synchrony but also on the correlation of the amplitudes for frequency f. To disregard amplitude and look only at phase synchrony the coherence can instead be calculated for signals of constant unit amplitude but the same phases as $\hat{x}_i(f)$ and $\hat{x}_j(f)$. This gives the quantity which in various equivalent forms has been called *phase locking value* (Lachaux et al., 1999) or *phase consistency* (Bruns, 2004). For a larger number of channels this generalizes to the phase synchronization order parameter introduced in Section 3.5.4. Measures based on asymmetry of the phase lag distribution have been proposed (Stam et al., 2007; Vinck et al., 2011) to disregard spurious zero-lag coherence due to volume conduction. Nonlinear cross-frequency coupling can also be determined for both phase (Jirsa and Müller, 2013; Yang et al., 2015) and amplitude (Bruns et al., 2000).

But as discussed in Section 1.1 brain oscillations and their synchrony show rapid transitions. They are not stationary. Hence for each of the measures above time-resolved versions can be defined, to analyze changes in synchrony between channels over time. To do this time-resolved cross-spectra are computed from the single-channel time-resolved complex spectra $s(f,t)$ introduced above (whether obtained by Fourier or wavelet methods), after which smoothing over time and/or frequency is applied (Mehrkanoon et al., 2011).

The connectivity statistics described above have in common that they are undirected. By contrast a family of methods including Granger causality (Seth et al., 2015) and transfer entropy (Chávez et al., 2003) give directed statistical measures of connectivity. These are applied to neuroimaging time series to attempt to characterize the direction of influence or information flow between brain regions.

The broad approach discussed thus far aims to infer dynamic relationships between brain regions from joint statistics of the time series alone. These measures are collectively called *functional connectivity*. The challenge is that each neuroimaging channel provides only a one-dimensional projection of the high-dimensional nonlinear activity of a brain region. Had we zero knowledge of the biophysical mechanisms underlying the time series, functional connectivity via descriptive statistics between channels would be the only feasible approach. However an increasing body of knowledge of microscopic and collective neural dynamics, neuroanatomy and semiology can be exploited as a second input to inform inference of brain interactions from time series. This is formalized in the Bayesian framework of dynamic causal modeling (Friston et al. (2003); Daunizeau et al. (2011) for a review). The directed interaction of brain regions inferred from time series using dynamical models of the underlying biology is called *effective connectivity*. The relationship between functional and effective connectivity is described by Friston (1994) and the approaches compared in detail by Valdes-Sosa et al. (2011). Dynamical models suitable for this purpose are a primary concern of this thesis and so they will be introduced further in Section 1.3.

1.2.2 Physiological origin of macroscopic neural oscillations

The physiological basis in changing cell membrane potentials of the electromagnetic field measured in neuroimaging was discussed above. But why do the collective fields at mesoscopic scale have pronounced oscillations at various frequencies?

At the single neuron level the dynamics of a wide variety of different neuron types have been individually studied and many of them can either exhibit self-sustained oscillations or noise-driven oscillations as a single cell when stimulated (Gerstner and Kistler, 2002; Izhikevich, 2007).

At the microscopic network scale, both negative feedback loops due to inhibitory interneurons, and synaptic and axonal propagation delays, can result in oscillations (Kopell et al., 2000; Campbell, 2007).

But at macroscopic scale, the dynamical mechanisms behind the rhythms with which clinicians are so familiar are not yet ascertained. Certainly what is measured is the summed, partly synchronous collective activity of between 10^3 and 10^5 neurons. But different possible mechanisms have been proposed and experiments have not yet been able decisively to resolve these competing hypotheses.

Even in the case of the most well studied oscillation, the approximately 10 Hz alpha rhythm, the origin is uncertain. Based on dynamical models some researchers hypothesize that the oscillations may arise because of the interaction of distal regions, specifically a ~ 100 ms round-trip axonal delay between neocortex and thalamus (Robinson et al., 2001). Whereas in other models self-sustained alpha-frequency oscillations arise locally in a cortical region, due to a combination of local inhibitory feedback and characteristic synaptic response times (Grimbert and Faugeras, 2006; Spiegler et al., 2010). This is consistent with experimental observation of a phase reversal at cortical layer V by Lopes da Silva and Storm Van Leeuwen (1977). Lopes da Silva et al. (1997) and Lopes da Silva et al. (1974) suggest that rather than self-sustained oscillations, scalp alpha rhythm may reflect noise filtered by frequency selective neural populations locally in both cortex and thalamus. Of course mechanisms at different scales are not mutually exclusive and if coexistent could interact between scales to reinforce or entrain each other. By comparing raw EEG and EEG enhanced by "dura" source reconstruction, Nunez et al. (2001) argue that the measured oscillations may involve multiple alpha rhythms, with the interaction of distributed and localized alpha oscillations.

For beta oscillations (13-30Hz) the picture is even less clear. The frequency of beta oscillations in some subjects shows a double peak in the power spectrum, possibly indicating two separate mechanisms (Jensen et al., 2005). Long distance coherence in the beta band between primary motor cortex and spinal neural populations has been associated with static preparation for movement (Kilner et al., 1999; Boonstra et al., 2009a). Robinson et al. suggest that beta oscillations arise as a 2:1 harmonic

of nonlinear alpha oscillations (Robinson et al., 2001). But studies also show that alpha and beta vary independently both in power and phase synchrony (Mehrkanoon et al., 2014) which is not possible if it is just a harmonic. It is possible that cortical coupling to subcortical areas may play a role, for instance the basal ganglia where beta-frequency oscillations are known to be generated by a system with local feedback (Terman et al., 2002).

The local network dynamics of interacting pyramidal and fast-spiking interneuron populations have been suggested as responsible for gamma band oscillations (Kopell et al., 2010). The generative mechanisms behind these frequency bands and the spatial scale at which they are generated require further investigation. These are questions for which computational models may prompt the design of experiments to test the possible mechanisms. Any successful explanation for their origin must explain the markedly different power of alpha and beta oscillations at different locations on the cortex, in terms of different neural populations and different local and distal connectivity.

1.3 Scales of biophysical modeling

We first review the most concrete mathematical modeling approaches. By *biophysical* we mean models whose parameters directly represent physiological quantities such as voltage or density, in contrast to abstract mathematical models to be considered in the next section.

As seen in the discussion of oscillations above, brain dynamics can be characterized on different spatial scales. Both biophysical and abstract phenomenological models have been used with success to improve understanding of behavior at each of these scales:

Single neuron models

This scale is accessible experimentally by microscopy and patch-clamp recordings of single neuron electrophysiology. While a canonical example is the model of Hodgkin and Huxley (1952), models for a single neuron can be complicated, with many coupled compartments capturing the computational properties of branching dendrites, for example (Herz et al., 2006).

Microscopic network models

By coupling single neuron models of different types in networks matching those found experimentally,

these models explore how computational function may be implemented locally at the microscopic scale (Brunel and Wang (2001), for example). Included at this scale are models of the encoding of information in patterns of discrete spikes (Gutkin et al., 2001; Averbeck et al., 2006).

Mesoscopic neural mass models

Discussed in detail below, these models attempt to capture the low-dimensional collective dynamics of a local region of brain tissue. This scale is comparable to that experimentally accessible by local field potential (LFP) recordings. These models are also building blocks for whole-brain generative models of EEG and MEG recordings (see *macroscopic models* below).

Neural field models

These model the local spatiotemporal propagation of neural activity and have been applied for example to model the epileptogenic process (Breakspear et al., 2006) and to understand pattern formation in neural tissue (Coombes, 2005). As seen in Jirsa and Haken (1996) and Robinson et al. (1997) the scale and level of abstraction is close to that of neural mass models. Neural field models may be most naturally written as integro-differential equations due to non-local axonal connectivity but reformulated as local partial differential equations (PDE) for analysis and simulation (Breakspear and Jirsa, 2007; Coombes, 2010). Notably, the spatial derivative is sometimes set to zero in the PDE to give a spatially uniform model in the form of an ODE (Breakspear et al., 2006; van Albada and Robinson, 2009), in that case effectively yielding a neural mass model, whilst retaining in that model a term that reflects the temporal effect of the spatially diffuse axonal connectivity. See also Daunizeau et al. (2009) for the explicit coupling of spatial modes to neural mass models, and Spiegler and Jirsa (2013) which examines discretization of a neural field by a network of neural mass models.

Macroscopic models of the whole brain

These models aim to understand connectivity and communication between brain regions as well as global network phenomena in the brain. Such models often take the form of a weighted, directed graph network structure, with a model of local dynamics at each node in the network, together with a model of the interaction between nodes, for example Honey et al. (2007), Deco et al. (2009), Petkov et al. (2014), Schmidt et al. (2014), Sanz-Leon et al. (2015), and Gollo et al. (2016). Nodes may be as simple as a phase oscillator (Section 1.4.1 below) or as complex as a neural mass model (the focus of the next section).

1.3.1 Neural mass models

In this thesis we are particularly interested in biophysical modeling at the mesoscopic and macroscopic scales, because that corresponds to the scale of oscillatory phenomena observable in neuroimaging.

At the mesoscopic scale, neural mass models aim to capture the collective dynamics of a small region of brain tissue. A single region being modeled may comprise 10^4 to 10^5 neurons, for which the model will represent a collective state of the local neural populations with as few as 2 to 20 collective degrees of freedom. Most commonly, neural mass models model a small area of neocortex, sometimes called a *cortical column* because it includes neural populations representing all layers of a cross-section of the folded neocortical sheet.[2] However neural mass modeling has also been applied to allocortical and subcortical brain regions including hippocampus (Wendling et al., 2002), thalamus (Lopes da Silva et al., 1974) and basal ganglia (van Albada and Robinson, 2009; Holgado et al., 2010).

The term *neural mass model* originates with the work of Freeman (1972, 1975) who used the abstraction *neural mass* for a local region of neural tissue and developed dynamical models based on experimental data to explore the potential for information processing at this collective, mesoscopic scale. Contemporaneously Wilson and Cowan (1972) developed two-dimensional dynamical models of interacting excitation and inhibition at this scale and other early contributors to the field include Nunez (1974) and Lopes da Silva et al. (1974). In the literature models of this type are sometimes called *mean field models* by reference to mean field methods in physics. For clarity we use *neural mass model* and reserve the term *neural field model* for systems of PDE or integro-differential equations modeling spatiotemporal evolution explicitly. In the case of neural mass models the collective variables most often used are the means of microscopic variables such as single-cell membrane potential or firing rate averaged across the population of each type of neuron. But this is not always the case, as discussed below.

Utility of neural mass models

Non-invasive neuroimaging, by EEG, MEG, EMG and fMRI are measuring collective activity at the mesoscopic and macroscopic scales. Modeling at this scale embodies the hypothesis that the generic

[2] Anatomically and functionally there is no organizational separation into discrete 'columns', except perhaps in specialized situations such as rodent whisker barrel cortex. The discretization is a modeling technique.

dynamical mechanisms responsible for large-scale phenomena of interest such as oscillatory transitions, connectivity biomarkers and seizure propagation are also at the same scale, and so can be captured by an effective model at this scale (Nunez, 1995).

This can be juxtaposed with an alternative approach of constructing large scale computational models of brain regions, while retaining microscopic detail of single-cell dynamics and individual connections (Markram, 2006). Such models necessarily contain millions of parameters and variables, leading to two difficulties. Firstly, without simulating the developmental process, the model can only be parameterized statistically, which loses the realism that was sought. Secondly, when brain-like phenomena are identified in a small region of configuration space, this does not then confer understanding. The complexity of such a model together with its non-linear dynamics makes this only marginally easier than apprehending the brain's mechanisms directly: through complexity the model loses explanatory power, as almost any dynamical behavior is present for some choice of parameters.

Beyond seeking to understand generative mechanisms of large-scale brain activity such as the interaction of oscillations between regions or the propagation of seizures, neural mass models are also used for a second purpose. As their parameters correspond directly to physiological quantities, and variables to measurable electrophysiological activity, the models can be inverted to estimate from experimental neuroimaging data the values of biological parameters that cannot be measured directly. This is done by Bayesian inference on the parameter space given the data (Friston et al., 2003; Daunizeau et al., 2011).

At the macroscopic scale, this approach can be extended to infer the structure and direction of macroscopic network interactions between brain regions, from experimental EEG or MEG data. Currently this is typically done by assuming that a particular neural mass model is accurate, then applying both Bayesian model selection and parameter estimation to a model space consisting of different candidate macroscopic networks of effective connectivity (David et al., 2005; Boly et al., 2011). The accuracy of the results depends in part on whether the neural mass model correctly captures local collective dynamics. One way to gain assurance would be to verify that the inferred results are robust to perturbation of the neural mass model parameters and to the substitution of comparable neural mass models. This can be incorporated in the Bayesian framework itself by extending both the model space and number of free model parameters, though there is a tradeoff with computational cost (Daunizeau et al., 2011).

Construction of neural mass models

Three broad approaches have been used in the literature to arrive at neural mass models. Firstly, phenomenological models have been developed based on experimentally measured mesoscopic behavior, such as the response to evoked potentials measured in local field potentials. This approach is exemplified by Freeman (1975), who used simultaneous measurement of EEG and single neuron spike trains to inform modeling of the collective voltage to firing rate gain as a sigmoid-shaped function and the impulse-response of a neural mass as a first or second order linear filter.

Lopes da Silva et al. (1974) used experimental data to parameterize a computational model at the microscopic network scale, then defined a neural mass model abstracting its collective behavior, in order to derive the oscillation dependence on parameters by linear analysis assuming a noise-perturbed steady state.

These were antecedents of the model of Jansen and Rit (1995) and models derived from it (Wendling et al., 2002; David and Friston, 2003; Zavaglia et al., 2006; Moran et al., 2007; Sotero et al., 2007; Spiegler et al., 2010; Babajani-Feremi and Soltanian-Zadeh, 2010; Goodfellow et al., 2011; Huang et al., 2011; Blenkinsop et al., 2012; Molaee-Ardekani et al., 2013; Vindiola et al., 2014; Kunze et al., 2016). Due to its wide use we take this Jansen-Rit model as a representative example to demonstrate various general methods in this thesis. The Jansen-Rit model will be described in detail in Chapter 2.

Other notable phenomenological models at this scale informed by mesoscopic experimental data include Wilson and Cowan (1972), Nunez (1974), Wright and Liley (1994), Jirsa and Haken (1996), Robinson et al. (1997), and Marten et al. (2009).

In the second approach, neural mass and neural field models are derived bottom-up from single neuron and microscopic network models. Such derivations often use assumptions of either zero correlation or perfect correlation in the states and inputs of constituent neurons within a region, or of small fluctuations around a common mean. A notable exception is Buice et al. (2010). The effect of variance of microscopic parameters and states within a population on the neural mass parameters may also be explicitly accounted for in such a derivation (Breakspear et al., 2003; Marreiros et al., 2008). Deco et al. (2008) derived the evolving population density $p(\mathbf{x}, t)$ from a leaky integrate-and-fire single neuron model by assuming all neurons have identical input distributions. From this a neural mass model could be obtained by replacing the population density for each neuron type with a point mass at the

mean state. Rodrigues et al. (2010) mapped parameters of a conductance-based single-neuron model to parameters of a neural mass model, discussing needed assumptions in detail. Zandt et al. (2014) derived parameters of neural mass firing rate sigmoids and synaptic impulse response curves from microscopic networks of Hodgkin-Huxley neuron models. Notably they included both first and second distribution moments as dynamical variables and demonstrated this gave a more accurate representation of the collective dynamics. Further examples of the bottom-up approach are given by Brunel and Wang (2001), Shriki et al. (2003), Cai et al. (2004), Wong and Wang (2006), and Dumont et al. (2014).

The third approach is *mode decomposition*, which uses as the population variables not mean values of single-cell variables, but the coefficients of a decomposition of the population density $p(\mathbf{x}, t)$ of a microscopic model. Often a diffusion approximation is assumed (Omurtag et al., 2000) so that the density evolution can be described by a Fokker-Planck equation (see Section 1.5.3 below). A decomposition may be chosen by either a discretization of the state space $\mathbf{x}$ (Omurtag et al., 2000), a spectral or singular value decomposition of the discretized Fokker-Planck operator (Knight et al., 1996), or by principal component analysis (PCA) of the output of the simulated microscopic model. They can also be defined dynamically, as a nonlinear function of the time-varying input to a region (Knight, 2000). Alternatively, modes may be chosen to have more direct physiological meaning as a partition of state space into regions with qualitatively different behaviors (Assisi et al., 2005; Stefanescu and Jirsa, 2008; Laing et al., 2010). Having defined collective variables as the coefficients of these modes, their evolution is then given by the action of the Fokker-Planck operator on each mode (Knight, 2000), or by projecting the microscopic equations of motion onto each mode (Assisi et al., 2005; Stefanescu and Jirsa, 2008).

1.4 Reduced oscillator models in neuroscience

The best mathematical model for a biological system depends on the questions being asked. It is often best to use the simplest model that embodies a hypothetical mechanism for the phenomenon of interest. This allows predictions to be made from the model that rest only on this mechanism. Besides being a basic principle in science, the preference for minimal models can also be understood in information theoretic terms (Baker, 2013).

In the case of models of oscillator synchronization, minimal models also make available a large body of existing mathematical theorems on synchronization phenomena, from phase transitions, clustering and chimera states to synchronization with time delays, nonlinear coupling and various network structures, existing results which can then be applied to the brain.

In this section we introduce minimal models of oscillations and oscillator transitions that aim to capture the generic mechanisms behind oscillatory network interactions that are observed in the brain. In contrast to the biophysical models of the previous section, parameters in the models can not always be interpreted directly with physical units and a physiological meaning.

1.4.1 Phase oscillator models

Where the phenomena being studied are the phase interactions of oscillations that do not change in power or frequency, one-dimensional phase oscillator models are used. In isolation each oscillator has the simple equation

$$\dot{\theta} = \omega, \tag{1.3}$$

with a phase angle θ advancing with a constant angular frequency ω (identifying states that differ by 2π). The coupling between these oscillators can be modeled in different ways, which is discussed further in Section 3.4.2.

As shown by Winfree (1967) this is a good approximation for a wide range of natural self-sustained oscillating systems in nature, specifically, where there is weak coupling and the state of each oscillator remains close to a linearly stable limit cycle (Kuramoto, 1984; Ermentrout, 1986).

Because the simplicity of the models makes many questions analytically tractable, there is an extensive literature of results on the synchronization of oscillations using coupled phase oscillators as a starting point. The most fundamental is the model of (Kuramoto, 1984) which showed that all-to-all coupling of the form

$$\dot{\theta}_j = \omega_j + \sum_{i=1}^{N} c \sin(\theta_i - \theta_j) \tag{1.4}$$

can be replaced by each oscillator coupling only to a mean field $\rho e^{\mathrm{i}\psi}$:

$$\dot{\theta_j} = \omega_j + c\rho\sin(\psi - \theta_j) \qquad \text{where } \rho e^{\mathrm{i}\psi} = \frac{1}{N}\sum_{i=1}^{N} e^{\mathrm{i}\theta_i}, \tag{1.5}$$

resulting in a system amenable to analysis. A transition to stable synchronization occurs dependent on coupling strength c and the distribution of natural frequencies ω_j. For the subsequent rich progression of results for the Kuramoto model see the review of Strogatz (2000).

Examples of application of one-dimensional phase oscillator models to synchrony in the nervous system include non-antisymmetric coupling between neural oscillations allowing collective oscillator death (Ermentrout and Kopell, 1990), pulsatile coupling causing stable phase locking (Ermentrout and Kopell, 1991), entrainment of a mesoscopic population by its input while individual neurons remain unsynchronized (Popovych and Tass, 2011), and experimental results that the functional network structures of epilepsy patients, inferred from EEG, cause global synchrony at weaker coupling than those of healthy controls (Schmidt et al., 2014, 2016).

Breakspear et al. (2010) provides an overview of the theoretical background of general one-dimensional phase oscillator models and a comprehensive review of the use of these models in neuroscience.

1.4.2 Phase-amplitude oscillator models

On the other hand, where the amplitude is either weakly stable or strongly perturbed by noise or coupling, a phase reduction is not possible. Oscillatory dynamics are then fundamentally two-dimensional with both phase and amplitude important to the interaction between oscillators. Two-dimensional phase-amplitude models are then appropriate.

An important special case is where nonlinear amplitude dynamics allow transition to and from the oscillatory regime, through either bifurcation or bistable switching. These are the minimal models most appropriate for studying the interaction of large scale oscillations in the brain, where empirical data shows power in each oscillatory band can change abruptly.

These models will be examined in detail in Chapter 3. This thesis is partly concerned with combining the strengths of the two modeling approaches — biophysical models with parameters that have direct

physiological meaning and oscillator models which are simple and tractable — by explicitly mapping the former to the latter while keeping physiological parameters intact.

1.5 Noise in computational models

In any high-dimensional dissipative system, fluctuations occur. In macroscopic models microscopic fluctuations are represented by noise processes. We first examine what kind of microscopic fluctuations are being modeled by noise in computational neuroscience models, then introduce their mathematical treatment.

1.5.1 Biophysiological features modeled by noise

Random fluctuations and stochastic processes are important in neuroscience at all scales from single cells to the whole brain, and these are represented differently in mathematical models at these different scales.

At the scale of single neurons, noise reflects random fluctuation in a wide range of microscopic physical processes in the cell, the most important being the random opening times of ion channels in cell membranes at all parts of the neuron: dendrites, soma and axons. This *channel noise* introduces stochasticity to many processes modeled in single neuron dynamics and also adds jitter to the timing of action potential propagation along axons (Faisal and Laughlin, 2007). Other stochastic processes at this scale include the timing of vessicle fusing in synapses, miniature postsynaptic currents, and fluctuating membrane resistance. Faisal et al. (2008) provides a comprehensive review.

At the microscopic network scale, the trains of discrete spikes carrying information between neurons in digital form can be represented mathematically as stochastic point processes. In the simplest case an input spike train to a cell assembly may be assumed to be a Poisson process with only the rate mattering. Alternatively correlations can be modeled in spike timing when investigating synchrony at the microscopic scale (Tchumatchenko et al., 2011). In many cases far from being detrimental the stochasticity serves important roles in optimizing information processing (Ermentrout et al., 2008; Faisal et al., 2008).

In phenomenological models of cognition and learning, the neural encoding of a quantity in the brain may be represented in an assembly of neurons by both the mean and variance of a spike firing rate across the population, where the variance encodes the precision or uncertainty.

At the mesoscopic scale, which is the focus of this thesis, the collective input firing rate from other brain regions is often modeled as a small white noise fluctuation about a mean value. A justification for this modeling is as follows. The total input to synapses of a single neuron in the target population represents discrete action potentials received from hundreds of neurons in the source population. For each level of excitation of the source population, making assumptions that each input is temporally uncorrelated and different inputs are independent, this may then be modeled as a Poisson process with a very high rate, which is then well approximated by Gaussian white noise (Brunel et al., 2001). Averaging this across cells of the target population, again assuming independence, the collective firing rate variance is reduced, while the high frequency content remains.

White noise differs greatly from power spectra of empirically measured time series at this scale (local field potential) but this is partly because what is measured corresponds to summed membrane potentials, *after* synaptic and dendritic filtering of the discrete point process input. The stronger objection to the white noise model is to the assumptions above of no temporal correlation and independence of different inputs. Modeling the input firing rate with white noise fluctuation serves to expose the response of the model to input across all frequencies.

The variables representing physiological inputs and outputs of a neural mass or neural field model are usually separated by the temporal filtering effect of synaptic and dendritic dynamics being modeled. This reduces the effect in a network model of high frequencies in the input. However it must be kept in mind when interpreting power spectra of simulated time series: depending on which projection of the state space to one dimension is chosen, it may be before or after this filtering in the model, thus giving spectra with very different physical meanings.

Randomly distributed parameters and states at the microscopic scale can also enter into neural mass models at the mesoscopic scale, via their distribution moments. For example Marreiros et al. (2008) discusses in depth the derivation of the neural mass sigmoid function from the distributions of neural state and membrane threshold in the neural population, with the static sigmoid shape embodying an assumption of static variance in neural state (see also Zandt et al. (2014)).

We now examine how this stochasticity is represented mathematically in neuroscience models at the mesoscopic and macroscopic scales.

1.5.2 Stochastic differential equations

Ordinary differential equations (ODE) of the form $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$ describe how system state variables $\mathbf{x}$ evolve in time, by specifying how their rates of change $\dot{\mathbf{x}}$ depend on the current state $\mathbf{x}$ and time t. They are used in many situations to model the low-dimensional dynamics of biological and other physical systems that change continuously in time.

A stochastic ordinary differential equation (SODE or just SDE) is a differential equation where the rate of change also depends on a random process $\mathbf{Y}(t)$:

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t; \mathbf{Y}(t)). \tag{1.6}$$

Here the driving process $\mathbf{Y}(t)$ may be any stochastic process. Given an initial condition $\mathbf{x}(t_0)$, $\mathbf{F}$ assigns to each realization $\mathbf{y}(t)$ of the process $\mathbf{Y}(t)$ a different solution $\mathbf{x}(t)$, called a *sample path*, which in general can depend on the prior history $\mathbf{y}(s) : t_0 \leq s \leq t$ (van Kampen, 2011).

We focus on the important special case most often used in neuroscience: where the physical system is assumed to be continuous and Markovian (change depending only on the present state), and with driving stochastic process the Wiener process $W(t)$, which is a standard diffusion for which increments $W(t + \Delta t) - W(t)$ are Gaussian distributed with mean zero and variance Δt, and where successive increments are independent, no matter how small and close in time. In other words this special case is useful for modeling continuous systems where the driving microscopic fluctuations have very short correlation times. The Gaussian assumption suits a wide range of physical diffusion processes because of the central limit theorem. This driving noise process is often called *Gaussian white noise* because its power spectrum has equal power at all frequencies (Gardiner, 2010).

We can conceptualize an SDE of this class as the limit of a discrete time process as the size of time steps Δt goes to zero, with diffusion within each time step (Jacobs, 2010). But there is more than one

way to do this: in one dimension the *Ito integral* can be defined informally[3] as

$$\int_0^t g(s)\,dW(s) = \lim_{N\to\infty} \sum_{n=0}^{N-1} g(n\Delta t)\Delta W_n, \tag{1.7}$$

where $\Delta t = t/N$ and $\Delta W_n = W((n+1)\Delta t) - W(n\Delta t)$.
Unlike a deterministic integral, it matters that we choose here to evaluate the diffusion coefficient function g at the start of the nth interval, not at the end. It can be shown that this form corresponds to an ideal regime where microscopic fluctuations have zero temporal correlation (van Kampen, 2011).

By contrast the *Stratonovich integral* can be informally defined as

$$\int_0^t g(s) \circ dW(s) = \lim_{N\to\infty} \sum_{n=0}^{N-1} \frac{g(n\Delta t) + g((n+1)\Delta t))}{2} \Delta W_n, \tag{1.8}$$

using the average of g evaluated at the start and end of the interval (Jacobs, 2010). It can be shown that this form naturally arises as the limit of a process with fluctuations having short correlation time τ_0 as $\tau_0 \to 0$ (van Kampen, 2011).

In the chapters that follow, the Ito system

$$x(t) = x_0 + \int_0^t f(x,s)\,ds + \int_0^t g(x,s)\,dW(s) \tag{1.9}$$

will be written in a more compact differential notation as

$$dx = f(x,t)\,dt + g(x,t)\,dW, \qquad x(0) = x_0. \tag{1.10}$$

Similarly the Stratonovich system

$$x(t) = x_0 + \int_0^t f(x,s)\,ds + \int_0^t g(x,s) \circ dW(s) \tag{1.11}$$

will be written in differential form with a $\circ$

$$dx = f(x,t)\,dt + g(x,t) \circ dW, \qquad x(0) = x_0. \tag{1.12}$$

[3] i.e. without discussing convergence of the limit.

The two limits (Ito and Stratonovich) seem similar. But using the same coefficient functions f and g, Ito and Stratonovich equations give two different processes with different paths resulting. However the notations describe the same class of processes: having correctly modeled a biological system with one notation, the equations can be converted between Ito and Stratonovich form by modifying f as shown in Section 4.4.2. (If g is constant then they are the same.)

It is important to note that when transforming an Ito equation (1.10) with a change of variables the rules of ordinary calculus do not apply. Instead there is a modified chain rule. (See Gardiner (2010)). But for a Stratonovich system (1.12), a change of variables satisfies the normal calculus chain rule. This is why in Chapter 4 we shall convert SDEs to Stratonovich form before applying transformations, then change them back to Ito form when needed. This avoids implementing separate Ito calculus transformations in the computer algebra system (Roberts, 2007).

If the noise coefficient function $g(x,t)$ is constant, e.g. $g(x,t) = \sigma$, this is called *additive noise*. Otherwise it is called *multiplicative noise*. In the neuroscience literature a system with additive noise,

$$dx = f(x,t)\,dt + \sigma\,dW, \tag{1.13}$$

is often written in *Langevin form*, appearing like an ODE but with some symbol such as $\xi(t)$, $\eta(t)$, ϵ or ϕ representing Gaussian white noise:

$$\dot{x} = f(x,t) + \sigma\xi(t). \tag{1.14}$$

Here $\xi(t)$, the "formal derivative" of the Wiener process $W(t)$ is not an ordinary function as $W(t)$ is not differentiable. Care must be taken if this notation is used with a non-constant noise coefficient $g(x,t)$ as it then must be clarified whether Ito, Stratonovich or another interpretation is intended for the multiplicative noise.

1.5.3 Fokker-Planck equations

For stochastic dynamical systems of the class discussed above, The Fokker-Planck equation (FPE) is a partial differential equation that describes the deterministic evolution of the probability density $p(\mathbf{x},t)$

of the state $\mathbf{x}$. A corresponding Fokker-Planck equation can be determined directly from the coefficients of the SDE when in Ito form (Risken, 1996). As the simplest example, given a one-dimensional Ito system,

$$dx = f(x,t)\,dt + g(x,t)\,dW, \tag{1.15}$$

the probability density $p(x,t)$ of the state evolves according to this Fokker-Planck equation:

$$\frac{\partial}{\partial t}\,p(x,t) = -\frac{\partial}{\partial x}(f(x,t)p(x,t)) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(g(x,t)^2 p(x,t)). \tag{1.16}$$

For brevity in the more complicated systems of Chapter 4 and Appendix C, we shall usually write this using an algebra of differential operators:

$$\partial_t\,p(x,t) = \left[-\partial_x f(x,t) + \frac{1}{2}\partial_x^2 g(x,t)^2\right] p(x,t). \tag{1.17}$$

For the n-dimensional case of deriving Fokker-Planck equations from Ito equations, see Risken (1996) or our example in Chapter 4. A particularly useful feature of the Fokker-Planck equation is that it is linear in $p(\mathbf{x},t)$, or stated another way, the Fokker-Planck operator (the right-hand-side of Eq. (1.17)) is a linear operator. This fact was crucial to the analysis and simulation of the population density models as discussed in Section 1.3.1.

For one-dimensional systems there is a one-to-one mapping between SDE and FPE. But for systems of dimension 2 or higher there is an infinite set of different SDEs corresponding to each FPE (Gardiner, 2010). These are all stochastically equivalent in the weak sense, meaning their solutions agree on the same evolving probability distribution for $\mathbf{x}$ at all times. But the temporal behavior of sample paths may differ between the two processes. Thus in dimension ≥ 2 there is a loss of information in passing from a SDE to a FPE.

However, the Fokker-Planck form allows us to bring to bear the methods of partial differential equations and linear maps on function spaces to understand the stochastic process.

1.5.4 Importance of correct numerical treatment of SDEs

When numerically simulating systems with noise it is necessary to use numerical integration algorithms that are specifically designed for SDE, otherwise the computed path will converge to the wrong solution. In the computational neuroscience literature this has not always been done correctly.

The most obvious difference is that the increment ΔW of a driving Wiener process at each integration time step needs to be scaled with the square root of the time interval, so that its expected variance $\langle(\Delta W)^2\rangle = \Delta t$. This is a basic defining property of $W(t)$ as described in Section 1.5.2. Naively using a noisy function with an ODE integration algorithm will in fact simulate a system with a different noise intensity. An adaptive step-size ODE algorithm will not only use the wrong mean noise intensity but also simulate a noise intensity that varies with the time step chosen by the algorithm, an error that can not be fully compensated by scaling the mean noise intensity. In particular, in a d-dimensional system, each time one of the variables goes close to zero the algorithm is likely to reduce the step size to preserve relative precision. This causes step size to vary systematically with system state, as illustrated in Figure 1.1. Thus even compensating for noise intensity a naive simulation may have resulting simulated dynamics that are qualitatively different from the true system.

In this thesis we implement stochastic numerical integration schemes to simulate the systems: in Chapter 2, the Stratonovich Heun algorithm (Rümelin, 1982; Burrage et al., 2004). In Chapters 5 and 6, the Stratonovich Runge-Kutta algorithm SRS2 of Rößler (2010). We compute repeated stochastic integrals by the methods of Kloeden and Platen (1992) and Wiktorsson (2001). Our implementation of these methods is described in detail in Appendix E.

1.5.5 Estimation of noise intensity for Jansen-Rit model

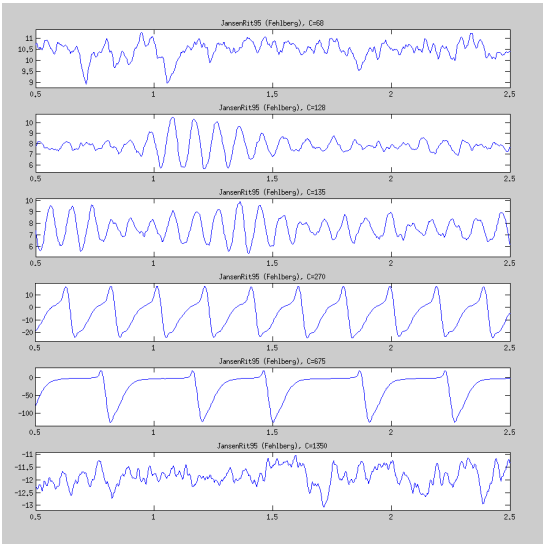
To simulate the Jansen-Rit model correctly in Chapters 2 and 6 of this thesis, we first need to determine here the actual noise intensity to use in these models to obtain their physiologically relevant behavior. Jansen and Rit (1995) and Wendling et al. (2000) used an adaptive-step-size ODE integration algorithm (Runge-Kutta). We shall see below that as a result the physiological level of firing rate fluctuations being simulated was misreported by a factor of roughly 30.

The noise used in the models of Jansen et al. (1993), Jansen and Rit (1995), and Wendling et al. (2000) is not fully characterized in those papers. But we can infer likely parameters and compare reproducing the results using both ODE and SDE integration algorithms. Jansen and Rit (1995) describes the noise as *"white noise [with] amplitude varying between 120 and 320 pulses per second"*. Assuming that a value of p was drawn uniformly from this range at each time step, due to the central limit theorem this is equivalent to using Gaussian white noise of the same standard deviation $\sigma_p = ((320 - 120)/2)/\sqrt{3} = 57.74\,\mathrm{s}^{-1}$, $\quad \langle p \rangle = 220\,\mathrm{s}^{-1}$.

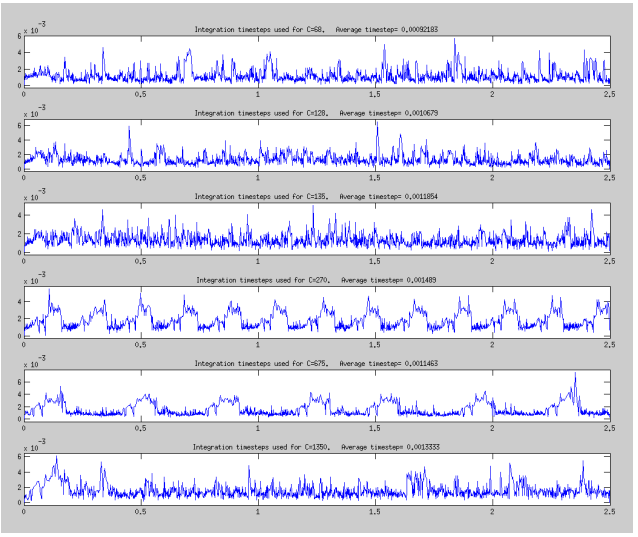
Given the bifurcation diagram of Grimbert and Faugeras (2006), this high level of variability in the input $p(t)$ was, prima facie, inconsistent with the transitions seen in Jansen and Rit (1995). To diagnose this issue we first used the Stratonovich-Heun SDE integration algorithm with the noise intensities specified by Jansen and Rit (1995) and Wendling et al. (2000). The results of those papers were not reproduced.

We then implemented the order 4-5 Runge-Kutta-Fehlberg ODE integration algorithm used by Jansen and Rit (1995) and Wendling et al. (2000) and simulated the models with that (incorrect) method, confirming their results are then reproduced in both cases.

To estimate a correction factor we next modified the RKF45 ODE integration algorithm to record its time step at each step. Using default parameters we estimate the distribution of step sizes chosen by the algorithm while reproducing the results of Jansen and Rit (1995) and Wendling et al. (2000). There was systematic variation with a mean time step of 0.0012 s.



A Results comparable to Fig. 3 of Jansen and Rit (1995)



B Time steps of RKF45 integration algorithm vary systematically with phase during simulation.

Figure 1.1: Simulation reproduces the results of Jansen and Rit (1995) incorrectly applying an ODE integration algorithm to a noisy system. The step size shows that the effective simulated noise strength spuriously varied with system state where no such feature exists in the model.

One can correct for noise intensity by multiplying by the square root of the estimated RKF45 mean step size of 0.0012 s and repeating Stratonovich-Heun SDE integration of the same system. (This corrected noise intensity is weaker by a factor of 28.9). With this correction a proper numerical integration of the SDE then replicates the qualitative findings of Jansen and Rit (1995) and Wendling et al. (2000). This is the correction factor that will be used in Chapter 2 and 6 simulations of the Jansen-Rit system. In this way we recover the actual effective noise intensity used by earlier authors and return to it the correct physiological interpretation as a standard deviation of the input firing rate.

1.6 Bifurcations

In any dynamical system a *bifurcation* is a transition between qualitatively different behaviors as a parameter is changed continuously.

More precisely in the specific case where the system can be expressed as an ODE system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{a})$ dependent on parameters $\mathbf{a}$, what we mean by "qualitatively different" is that as $\mathbf{a}$ is moved continuously past a certain point, the vector field flow defined by $\mathbf{f}$ is *topologically nonequivalent* before and after that point, in the sense that there is no continuous mapping between them with continuous inverse

that preserves orbits following the flow (Kuznetsov, 2010). As discussed in Section 1.1, bifurcations are one of the several classes of dynamical transitions that potentially allow flexible changes to occur in the large scale dynamics in the brain.

Evidence for mesoscopic dynamics near instability

In contrast to modeling at the microscopic scale, where the range of dynamics of healthy neurons is known to include nonlinear behavior such as limit cycles, modeling at the larger scale of mesoscopic neural masses or neural fields often assumes that the dynamics at this scale operate close to a stable fixed point where input fluctuations result in only small and brief perturbations of the population state. As introduced in Section 1.1, there are theoretical reasons to consider that collective neural activity near linear instability is adaptive for brain function, including facilitating the observed rapid switching of oscillation power at different frequencies (Boonstra, 2007). Theoretical models also show the connection of criticality to optimal information processing (Friston, 2000; Friston et al., 2012; Kinouchi and Copelli, 2006; Shew and Plenz, 2013; Gollo and Breakspear, 2014; Chialvo, 2014).

Empirical evidence for activity close to instability at the mesoscopic scale includes long autocorrelation times of amplitude fluctuations of human brain oscillations in EEG, with power law decay (Linkenkaer-Hansen et al., 2001), bistable switching of alpha oscillation between two distinct modes (Freyer et al., 2009), increased variance and critical slowing (Kelso et al., 1986; Scholz et al., 1987), scale-free cortical activity measured by ECoG (Bedard et al., 2006; He et al., 2010), intermittent nonlinear structure within (Stam et al., 1999) and between (Breakspear and Terry, 2002) surface EEG channels, and spatiotemporal correlations with power law statistics in intracranial depth recordings (Priesemann et al., 2013). Such indicators will be discussed further in Chapter 2.

Local bifurcations of equilibria

While many bifurcations are global, affecting the structure of the flow $\mathbf{f}$ across an extended region of the state space, the methods in Chapters 3 and 4 specifically apply to various types of *local bifurcations* of equilibrium points. For an ODE system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, an *equilibrium point* is a location $\mathbf{x}_c$ in the state space where $\mathbf{f}(\mathbf{x}_c) = 0$, so that point remains invariant under the flow defined by $\mathbf{f}$. An equilibrium point $\mathbf{x}_c$ is *asymptotically stable* if for starting points close to $\mathbf{x}_c$, the state $\mathbf{x}(t)$ is drawn arbitrarily close to $\mathbf{x}_c$ as time $t \to \infty$. To assess asymptotic stability of an equilibrium point we first check the stronger property of *linear stability* by evaluating the Jacobian matrix $D\mathbf{f}(\mathbf{x})$ at the equilibrium

point. For real $\mathbf{f}$ the eigenvalues of the Jacobian are real or occur in complex conjugate pairs. If its eigenvalues all have negative real parts then the equilibrium point is asymptotically stable. Stronger than this, it is *linearly stable*: if the state $\mathbf{x}$ starts close to the point $\mathbf{x}_c$, it will collapse exponentially towards $\mathbf{x}_c$ like the linear equation $\dot{x} = -ax$.

Now consider a system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \epsilon)$ depending on a parameter ϵ, with stable equilibrium point $\mathbf{x}_c(\epsilon)$. As ϵ is changed continuously, the eigenvalues of the Jacobian at the equilibrium also change continuously, and some may cross the imaginary axis, causing linear stability of the equilibrium to be lost. This is a *local bifurcation* of the equilibrium point, as the flow is qualitatively changed in a local neighborhood of the equilibrium.

1.6.1 Increased importance of noise near bifurcation

Close to bifurcation the effect of noise perturbation is greatly increased (Haken, 1983) and this results in characteristic changes in time series statistics (Scheffer et al., 2009).

When close to stable equilibrium, a system's local behavior is approximately linear, and in analyzing experimental time series near a stable equilibrium the theory of linear time-invariant systems applies. In this case any perturbations from the equilibrium (including noise driven perturbations) are quickly dissipated and decay exponentially in time back to the stable state. Thus only response behaviors on short time scales are observable. Because response amplitude is damped the response is also more easily masked by measurement noise (such as the activity of other brain regions). In this stable regime the best connection that can be made between the model and observed time series is to determine from the model a transfer function relating frequencies in the input perturbation to frequencies in the resulting system output (e.g. Robinson et al., 2002; Robinson et al., 2004).

Similarly when the system is in a regime of stable oscillations on a limit cycle, perturbations away from this attractor also decay exponentially back to the limit cycle. If a one-dimensional time series can be observed in experiment, the dominant behavior is the oscillation at the base frequency. Beyond this the relative strength of harmonics (integer multiples of the oscillator base frequency) in the output may distinguish between different classes of nonlinear oscillator (Chatterjee, 2009).

But it is close to the bifurcation between these two regimes, where linear stability is temporarily lost, that the susceptibility of the system to intrinsic or extrinsic perturbations is greatest. As the response is not so quickly dissipated the real system is thus likely to divulge more information about the underlying dynamics. Here behaviors on longer time scales can potentially be observed. Where models explain a transition of behavior in the real system, and it is possible experimentally to observe the system through this transition, this offers the best opportunity to measure the statistical behavior of the real system and compare to models.

1.6.2 Andronov-Hopf bifurcations

To exemplify the methods we focus on the supercritical Andronov-Hopf bifurcation because it is the simplest and perhaps the most common mechanism by which oscillations start and stop. In short this is a transition in which a stable equilibrium changes into a stable cycle of oscillations as a parameter is changed. This situation was first studied by Poincaré (1892) then by Andronov and Witt (1930) and Hopf (1942). A thorough treatment is provided in Kuznetsov (2010) and Marsden and McCracken (1976). Here we present a brief and informal overview.

Referring to the system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \epsilon)$ of the previous section, if the equilibrium point $\mathbf{x}_c(\epsilon)$ undergoes a local bifurcation where a single conjugate pair of nonzero eigenvalues crosses the imaginary axis as ϵ is changed (having values $\pm \mathrm{i}b$ when $\epsilon = 0$, say) then this is an *Andronov-Hopf bifurcation* (Hopf bifurcation for short). But whether stable oscillations emerge locally at this point depends on whether the Hopf bifurcation is *supercritical*, *subcritical* or *degenerate*. This is determined by the *first Lyapunov coefficient* $l_1(0)$, a real value dependent on the coefficients of the Taylor expansion of the flow $\mathbf{f}(\mathbf{x}, \epsilon)$ at $\mathbf{x}_c$ up to third order together with an adjoint pair of eigenvectors corresponding to the crossing eigenvalues. (For the detailed formula in n dimensions see Kuznetsov (2010)). In fact the value of $l_1(0)$ depends on the convention chosen for scaling the eigenvector used in its calculation: it is the sign of $l_1(0)$ that matters.

If $l_1(0) > 0$ the Hopf bifurcation is *subcritical*: the stable equilibrium merges with an unstable limit cycle at the bifurcation point and becomes an unstable equilibrium. If $l_1(0) < 0$ then it is *supercritical*: at the bifurcation point the stable equilibrium becomes a stable limit cycle, locally attracting the state $\mathbf{x}(t)$ to orbit with self-sustained oscillations of frequency $\approx 2\pi/b$, surrounding an unstable

equilibrium point. As the bifurcation parameter ϵ continues past the bifurcation point $\epsilon = 0$ the small oscillations will initially increase with amplitude proportional to $\sqrt{\epsilon}$. If $l_1(0) = 0$ the Hopf bifurcation is *degenerate*. This case will be discussed further below, as it is also useful to neuroscience modeling.

For the supercritical Hopf, it is important to emphasize that locally the system stays asymptotically stable throughout: before the bifurcation a nearby state collapses exponentially to the equilibrium point; after the bifurcation a nearby state collapses exponentially to the limit cycle. But close to the bifurcation the system loses *linear* stability. It is still asymptotically stable, but the stability is weaker and the collapse to the attracting set is no longer exponential in time. This will be important for the statistics of time series in later chapters.

Figure 1.2 illustrates the simplest possible cases: Hopf bifurcations with circular symmetry occurring in two-dimensional systems on a flat plane as a bifurcation parameter β is increased.

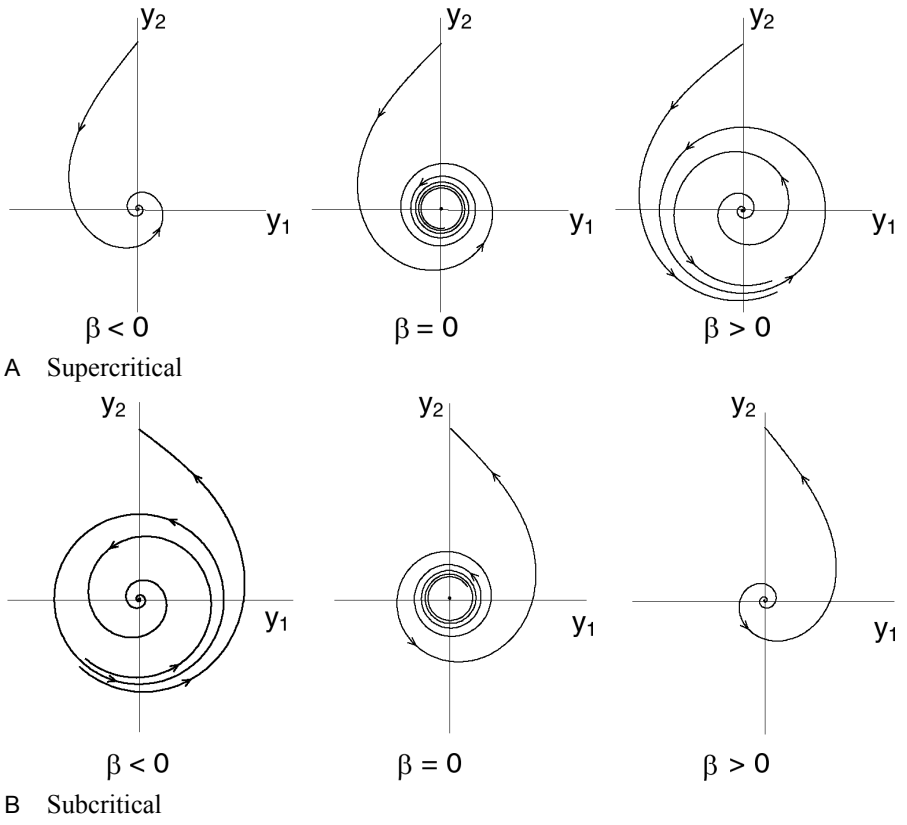


Figure 1.2: Supercritical and subcritical Andronov-Hopf bifurcations in two dimensions.
Figure by Yuri A. Kuznetsov, Creative Commons Attribution-NonCommercial-ShareAlike license.

Hopf bifurcations in biological systems are usually not as simple as those in the figure above. Firstly in two dimensions the oscillations are in general not circular in the state space. Secondly in an n-dimensional system the Center Manifold Theorem guarantees that the transition to oscillation will be qualitatively similar to the two dimensional picture: there exists a two-dimensional surface on which the limit cycle will emerge (Guckenheimer and Holmes, 1985). However that surface may be curved in arbitrary ways in the n-dimensional space in which it is embedded, so that the oscillations are not constrained to a plane. Some figures with examples of this appear in Chapters 3, 5 and 6.

Both supercritical and subcritical Hopf bifurcations occur in many contexts and different scales in neuroscience, from the cell membrane model of Hodgkin and Huxley (Troy and McLeod, 1976) to the Jansen-Rit neural mass model (Grimbert and Faugeras, 2006).

Besides systems of the form considered above, Hopf bifurcations often occur in systems with explicit time delays (modeled with delay differential equations (DDE) rather than ODE). With a delay, one state space dimension is then sufficient for a Hopf bifurcation, rather than two (Hale, 1971), and this too is a very common generic mechanism for transitions to oscillations in biology (for an example in the context of a neural field model see Roberts and Robinson (2012)).

Hopf bifurcations can similarly cause a transition to temporal oscillation at each point in a continuous medium (Marsden and McCracken, 1976), such as a two-dimensional sheet of cortical tissue modeled by PDE, and in conjunction with a Turing bifurcation at the same point can give rise to spatiotemporal oscillation – traveling waves (Coombes, 2010).

Generalized Hopf bifurcation

If the Hopf bifurcation is degenerate, with $l_1(0) = 0$, the flow near $\mathbf{x}_c$ will have one of several possible qualitatively different structures. The possibilities can be enumerated by an *unfolding* of the bifurcation (Golubitsky et al., 1985) and to distinguish which one is present in a particular system requires the fifth and higher order terms of the Taylor expansion of $\mathbf{f}$. The most important case for neuroscience is a two-parameter bifurcation at that point called a Generalized Hopf or Bautin bifurcation (Kuznetsov, 2010). This is a point where a subcritical Hopf becomes a supercritical Hopf as a second parameter is varied. In computational neuroscience modeling, a branch of subcritical Hopf bifurcations may be used close to the point of a Generalized Hopf bifurcation so that there is a larger stable limit cycle surrounding the bifurcating equilibrium allowing bistability between these two attractors (Kalitzin et al.,

2010; Benjamin et al., 2012; Petkov et al., 2014). This is the situation shown in region ③ of Figure 1.3 below. In this figure the point at the origin is the Generalized Hopf, but bistability emerges at an ordinary subcritical Hopf bifurcation as the system crosses the green line from region ② to ③.

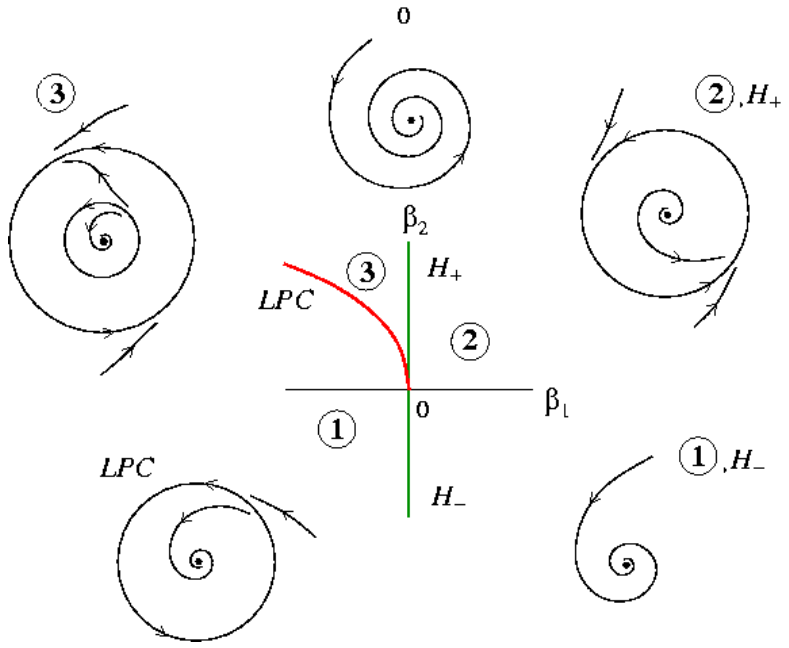


Figure 1.3: Generalized Hopf (Bautin) bifurcation. Here a Hopf bifurcation occurs when crossing the vertical axis: it is subcritical at H_+ and supercritical at H_-.
Figure by Yuri A. Kuznetsov, Creative Commons Attribution-NonCommercial-ShareAlike license.

1.7 Analytical methods and simulation

This thesis makes use of both large scale computer simulation of models and analytical treatment of models by asymptotic approximation and algebraic transformation. These tools have complementary roles.

The behavior of systems, even with low dimension $d \geq 3$ can be complex. Numerical simulation of mathematical models enables experiment on their dynamics, can prompt new hypotheses for testing on the real biological systems, and sometimes suggests properties that can then be confirmed mathematically or empirically to be true more generally. Bounds can be put on parameters by experiment and biophysical reasoning, to focus the target of simulation.

The important limitation is that each individual simulation run only gives answers for a single choice of parameters. Functional relationships can sometimes be inferred by large scale multiple simulations of a lattice of points in a region of parameter space. But especially for stochastic behavior with a sufficiently large and high dimensional parameter space, that becomes infeasible even with high performance computing resources. Bifurcations or behaviors that occur on a boundary or region of small measure may also be missed when sampling parameter space. Thus there need to be methods to know where to look.

Mathematical analysis of a model can sometimes demonstrate properties in general for a range of model space and parameter space. It may mark out the boundaries between regions of different behavior, to narrowly focus computer simulations. Most importantly, by deriving functional relationships rather than quantities, analytical methods can sometimes directly uncover the generic mechanism behind the observed behavior. Because simpler systems are often more analytically tractable, approximation and asymptotic methods play a role in the analytical approach. This is exemplified in the chapters that follow, which combine computer simulation with analytical transformations and approximation.

1.7.1 Asymptotic approximations

The methods we present for analyzing local bifurcations of equilibria in a model are approximations valid only in a neighborhood of the bifurcation point in phase space and parameter space. For example an early step is to approximate the flow $\mathbf{f}$ close to the bifurcation point with a truncated Taylor series approximation (that is, the sum of the first n terms). When obtaining a weak approximation of a stochastic system by averaging in Chapter 4 we also use more general power series approximations. Finite asymptotic series were first made rigorous by Poincaré (1886). For the definitions of pointwise and uniform asymptotic approximations and expansions, see Murdock (1999).

We must distinguish between the neighborhood of convergence of an infinite series and the neighborhood of useful applicability of a truncated series approximation. *Neither implies the other*. Depending on coefficients, the initial terms of a series expansion may be a poor approximation even though the series converges. Conversely, even when an asymptotic series is nowhere convergent, its truncation may provide a useful approximation. It is the size of the error of the partial sum that matters, and

whether that error is small enough for practical use. The convergence of the tail sum is not relevant to asymptotic approximation (Murdock, 1999). Thus the size of the region near the bifurcation point in which our approximations remain useful will depend on both the system and the intended use.

When we deal with systems perturbed by noise, in Chapters 4-6, the approximations are also valid only for sufficiently weak noise. For sufficiently high noise intensity σ the noise will push the system state away from the region of phase space where the series approximations are valid. In particular, when systems are multistable, a noise process with unbounded amplitude distribution (such as our Gaussian white noise) will push the state away from the attractor of interest into the basin of attraction of another attractor with a finite expected switching time. Characterizing the behavior of one attractor, such as a limit cycle of oscillations, in isolation is only meaningful where weak noise means this switching time is long compared to the behavior of interest.

In this context large scale numerical simulations are a useful check that our truncated series and other approximations preserve the properties of interest at the noise levels biologically specified in the model. We use this in Chapter 6 as a check that the stochastic weak model approximately preserves specific oscillation statistics.

Considerations for asymptotic scaling near a bifurcation in the presence of noise are discussed in more detail in Section 4.4.1. A criterion balancing noise strength and stability is given in Section 4.4.3.

1.7.2 Normal form transformations

Intuitively a normal form is a "simpler" dynamical system having "qualitatively the same" local flow in the state space as the original system.

For example for a dynamical system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ if all eigenvalues of the Jacobian $A = D\mathbf{f}(\mathbf{x})$ have non-zero real part then the normal form is a simple linear system: removing all but the linear term of $\mathbf{f}$ gives its equation: $\dot{\mathbf{x}} = A\mathbf{x}$. For each type of local bifurcation of equilibria there is a simple normal form system of minimal dimension exhibiting that bifurcation at the origin. At a Hopf bifurcation linear and cubic terms will be present in a normal form, with fifth or higher order terms present if the Hopf bifurcation is degenerate.

Normal form theory includes methods to find explicitly a smooth mapping from the original system to a simplified normal form.

Deterministic normal form tranformations

To be precise, a "simpler" system here means firstly one that is more symmetrical, in the sense that the flow vector field $\mathbf{f}$ is unchanged (equivariant) under the action of symmetry groups, and secondly where the bifurcation occurs in a decoupled subspace of minimal dimension: its center eigenspace is a center manifold of the bifurcation. Transforming to a more symmetrical system removes non-symmetrical terms from the mathematical expression for $\mathbf{f}$, so the equations of the normal form system are also simplified. And because the subspace containing the bifurcation is decoupled from other variables, it can be treated separately, omitting the other equations: a center manifold reduction. Provided the other eigenspaces are all stable, the state of the complete system will collapse exponentially to this subspace so that after an initial transient its behavior is given by the simple, decoupled system of minimal dimension. This is the reduced normal form.

But there can be different definitions for "qualitatively the same" in the discussion above, leading to different definitions of a normal form. Each involves partitioning the possible vector fields $\mathbf{f}$ according to some equivalence relation.

The definition with the broadest equivalence classes is that of *topological normal forms* (Kuznetsov, 2010). By this definition two systems are equivalent if there is a homeomorphism (a continuous map with continuous inverse) that maps one to the other preserving orbits along the flow.

By contrast for Poincaré normal forms (also called equivariant normal forms) which includes the semisimple and inner-product normal form styles of Murdock (2003), the system is already in normal form when the higher order terms have the symmetry (equivariance) of the linear term. In mathematical terms this means the system is in normal form when its linear term is in Jordan canonical form and its linear term and nonlinear terms commute as vector fields:

$$[\mathbf{f}(\mathbf{x}),\ A\mathbf{x}] \ = \ 0, \tag{1.18}$$

where $[\ ,\ ]$ is a Lie bracket of vector fields defined by

$$[\mathbf{f}(\mathbf{x}),\ \mathbf{g}(\mathbf{x})] \ = \ (D\mathbf{f}(\mathbf{x}))\mathbf{g}(\mathbf{x}) - (D\mathbf{g}(\mathbf{x}))\mathbf{f}(\mathbf{x}) \tag{1.19}$$

(Murdock, 2003). Computing the Poincaré normal form is equivalent to a linear projection of the flow vector field $\mathbf{f}(\mathbf{x})$ into a chosen subspace of fields that satisfy this symmetry property. In this way, the method makes the minimal changes required to the system to achieve the symmetry, and thus has finer equivalence classes: those systems that project to the same image are equivalent. Section 3.3 will introduce methods of deriving a Poincaré normal form transformation. This will then be automated in a Mathematica package that can be applied to an arbitrary system near any local bifurcations of equilibria, described in Appendix D.

The topological normal form is obtained from the Poincaré normal form by a further nonlinear rescaling of time t and of the bifurcation parameters. For details of this see Kuznetsov (2010).

In application we shall see that the different kinds of normal form have different roles. For each bifurcation type its topological normal form has maximal simplicity. By transforming a biophysical model to its Poincaré normal form rather than topological normal form more properties are quantitatively preserved while achieving the symmetry and decoupling, which we shall see is important for preserving the response to perturbation if the system is not isolated.

Stochastic normal form transformations

There are also developments of normal form theory specifically to treat transformation of stochastic dynamical systems (Arnold and Imkeller, 1998; Roberts, 2008). These result in stochastic normal forms with symmetrical noise terms, an objective we instead achieve in Section 4.4.4 by averaging a Fokker-Planck operator, at the cost of mapping to a weak approximation.

The approach of Roberts (2008) aims to separate slow and fast processes to the greatest extent possible in a stochastic system, which then allows highly efficient numerical simulation (that is, with a longer time step). To achieve this some stochasticity is embodied in the mapping between the original system and its normal form, which becomes a random transformation, while some stochasticity remains in the reduced normal form system. To achieve this aim of efficient sample path simulation, it is also important to use strong models: those whose sample paths directly correspond to sample paths in the original system.

These aims differ in some ways from those of the present work, where we prefer all stochasticity to be retained in the simplified system so that the total effect of a noise perturbation can be juxtaposed with the deterministic terms. As the main focus is the effect of biological parameters on oscillation time

series statistics, weak models that preserve those statistics can also be considered. Therefore in this thesis we explore the idea of viewing stochastic dynamical systems in the coordinate systems defined by their deterministic normal forms and see how far this can be taken. In Chapters 4-6 this proves to be a fruitful approach for answering our questions about stochastic oscillations in biophysical models, and one that may also have broad application beyond neuroscience.

1.8 Structure of this thesis

The thesis is structured as follows:

Neuroscience simulations in Chapter 2 identify sometimes counterintuitive effects of noise near the onset of oscillations in a model of cortical tissue. This motivates the search for a way to analyze stochastic systems near bifurcation to explain or predict these effects.

Chapter 3 introduces normal form transformations for deterministic systems. We show how this can be applied in neuroscience to constrain the parameters of reduced models, and demonstrate that this is useful for computational studies of network synchrony near bifurcation.

Chapter 4 presents a new method for applying normal form transformations to systems with noise. This method allows analytically to predict from a biophysical model how stochastic oscillation phase and amplitude dynamics near a Hopf bifurcation will change as biological parameters and noise parameters are varied.

Chapter 5 examines which statistical properties of oscillating time series are preserved by the transformations of Chapter 4.

The new method is then applied for real in Chapter 6, analyzing the Jansen-Rit neural mass model to answer the questions raised in Chapter 2.

Chapter 2

Critical fluctuations in cortical models near instability

Computational studies often proceed from the premise that cortical dynamics operate in a linearly stable domain, where fluctuations dissipate quickly and show only short memory. Studies of human electroencephalography (EEG), however, have shown significant autocorrelation at time lags on the scale of minutes, indicating the need to consider regimes where non-linearities influence the dynamics. Statistical properties such as increased autocorrelation length, increased variance, power law scaling, and bistable switching have been suggested as generic indicators of the approach to bifurcation in non-linear dynamical systems. We study temporal fluctuations in a widely-employed computational model (the Jansen-Rit model) of cortical activity, examining the statistical signatures that accompany bifurcations. Approaching supercritical Hopf bifurcations through tuning of the background excitatory input, we find a dramatic increase in the autocorrelation length that depends sensitively on the direction in phase space of the input fluctuations and hence on which neuronal subpopulation is stochastically perturbed. Similar dependence on the input direction is found in the distribution of fluctuation size and duration, which show power law scaling that extends over four orders of magnitude at the Hopf bifurcation. We conjecture that the alignment in phase space between the input noise vector and the center manifold of the Hopf bifurcation is directly linked to these changes. These results are consistent with the possibility of statistical indicators of linear instability being detectable in real EEG time series. However, even in a simple cortical model, we find that these indicators may not necessarily be visible

even when bifurcations are present because their expression can depend sensitively on the neuronal pathway of incoming fluctuations.

2.1 Background

Computational models of neocortex and other brain structures have proved very useful for a range of research problems in neuroscience (Braun and Mattia, 2010; Friston and Dolan, 2010). Interpreting empirical data using dynamical models is particularly fruitful in neuroimaging, where underlying processes are obscured by the low temporal resolution of fMRI or the coarse spatial source resolution of EEG/MEG. This allows testing of hypotheses about internal dynamical mechanisms (e.g. Freyer et al., 2012) and, through model inversion, the estimation of neural and connectivity parameters that cannot be observed directly (Friston et al., 2003). In contrast to modeling at the microscopic scale, where the range of dynamics of healthy neurons is known to include nonlinear behavior such as limit cycles, modeling at the larger scale of mesoscopic neural masses or neural fields often assumes that the dynamics at this scale operate close to a stable fixed point where input fluctuations result in only small and brief perturbations of the population state. This premise is predicated on the diffusion approximation that states that correlations amongst neuronal inputs are reduced as the size of the population increases (for review, see Deco et al., 2008). This approach enables the calculation of spectra from the composition of transfer functions, a powerful technique that allows physiological parameters to be estimated from non-invasive functional neuroimaging (Friston et al., 2003) and neurophysiological (van Albada et al., 2010) data.

Dynamic instabilities in models at the larger scale of neural masses have typically been associated with the pathological activity of epileptic seizures (Robinson et al., 2002; Breakspear et al., 2006; Wendling et al., 2000). However, empirical data shows that such instabilities may also underlie healthy neural activity (Freyer et al., 2009, 2011). Indeed, the Jansen-Rit neural mass model (Jansen and Rit, 1995) and its derivatives (Wendling et al., 2002; David and Friston, 2003; Moran et al., 2007; Zavaglia et al., 2006; Sotero et al., 2007; Spiegler et al., 2010) reach bifurcations where fixed points become linearly unstable while still within the healthy physiological range of parameters. In fact, oscillations in the model output that have been identified with normal cortical alpha activity have been shown to arise

from limit cycle activity following a supercritical Hopf bifurcation (Grimbert and Faugeras, 2006; Spiegler et al., 2010).

The term "linear instability" here does not necessarily imply that the dynamics of the system as a whole lose stability. Indeed, in the case of the supercritical Hopf bifurcation, stability of the attractor is maintained as it deforms continuously from a stable fixed point to a stable limit cycle, which then increases in size in the phase space. Hence, there is no discontinuous transition. The distinction is that the dominant dynamics in the system are no longer linear. The presence of quadratic and higher order flow terms that become significant in the neighborhood of a bifurcating fixed point have a profound influence on the system's statistical properties and its response to stochastic perturbations.

The putative presence of linear instabilities in healthy, mesoscopic cortical activity is ultimately an empirical question that must be answered with reference to the theory of nonlinear stochastic dynamical systems. For a wide range of systems, statistical measures such as increased autocorrelation, increased variance and bistable switching have been proposed as generic indicators that the system is losing linear stability on approaching a bifurcation (Kelso, 2010; Scheffer et al., 2009). Increased autocorrelation length is a direct consequence of critical slowing-down, which occurs as the strength of attraction to a stable fixed point becomes weaker before changing to equally weak repulsion. Long-range correlations may also reveal a transition from exponential to power law relaxation in the vicinity of linear instabilities as a result of the higher order (nonlinear) flow terms.

Within neuroscience, statistical indicators of bifurcations have been studied at a range of scales, in both computational models and empirical analyses. In the context of single neuron models, increase of variance close to a bifurcation and the spectral peak near a Hopf bifurcation have been examined (Steyn-Ross et al., 2006). Spectral features and variance close to instability have been explored in large-scale mean field corticothalamic models (Robinson et al., 1997, 2002; Roberts and Robinson, 2012) and mean field models of the brainstem and hypothalamus (Robinson et al., 2010). Slowing down, instability and bifurcations have also been studied at the highest level of brain function, particularly in human movement. For example, increased variance and critical slowing have been observed in human bimanual motor control (Kelso et al., 1986; Scholz et al., 1987) and are explained by a simplified phenomenological model of coordination (Haken et al., 1985).

In addition to the analyses of empirical data contained within these computational studies, signatures of transitions in neuroimaging data have been the subject of a number of predominantly empirical studies. Amplitude fluctuations of human brain oscillations have been shown to have long time autocorrelations with power law decay in electroencephalography (EEG) (Linkenkaer-Hansen et al., 2001), consistent with effects expected near linear instability. Scale-free cortical activity has also been reported in surface electrocorticogram (ECoG) activity, although the significance, scaling coefficient and likely mechanisms remain contested (Bedard et al., 2006; He et al., 2010; Miller et al., 2009). Similarly, Stam and de Bruin (2004) reported scale-free fluctuations in the degree of synchronization between surface EEG recordings. These findings are consistent with prior reports of intermittent nonlinear structure within (Stam et al., 1999) and between (Breakspear and Terry, 2002) surface EEG channels. More recently, Freyer et al. (2009) found that 10 Hz oscillations showed intermittent switching between two distinct bistable modes, although the dwell times within each mode followed a stretched exponential, not a power law decay.

The objectives of the present study are to examine linear instabilities in the Jansen-Rit model, a closed set of equations describing the activity of a small cortical region and one of the simplest cortical neural mass models. At the same time it is a base upon which many extensions and derivative models have been built (Wendling et al., 2002; David and Friston, 2003; Moran et al., 2007; Zavaglia et al., 2006; Sotero et al., 2007; Spiegler et al., 2010). The phenomena which we report in this simple model therefore highlight the possibility of similar behavior in a wider class of models. We focus on one key indicator of linear instability (autocorrelation length) and one important bifurcation (supercritical Hopf). Time series for each neural population in the model are generated for sets of parameters approaching a bifurcation. We then test whether the autocorrelation indicator of proximity to bifurcation is reliably detectable in the time series of the pyramidal population and also examine scaling properties of fluctuations in this time series. In this way we explore whether simple bifurcations at the population scale have the potential to contribute to indicators such as lengthened autocorrelation times and power law scaling of fluctuations reported in human EEG data.

2.2 Models, simulation and analysis methods

2.2.1 Jansen-Rit neural mass model

Building on the earlier work of Lopes da Silva et al. (1974) and Wilson and Cowan (1972), Jansen and Rit developed a simple computational model of a small cortical region (Jansen et al., 1993; Jansen and Rit, 1995). The model produces an output signal similar to spontaneous EEG alpha oscillations, and also shows responses similar to evoked potentials following pulsatile input. The Jansen-Rit model is a closed set of differential equations that describe the local average states of three interconnected neural populations (Figure 2.1), excitatory interneurons, pyramidal cells and inhibitory interneurons. Here we follow David and Friston (2003) in identifying the excitatory interneurons in the model with layer IV spiny stellate cells. The spiny stellate and pyramidal neurons are both excitatory and both populations receive external input, although only the pyramidal cells project out of the local region.

Each second order equation in the model corresponds to a population of synapses and their post-synaptic dendritic processes (Freeman, 1992; Deco et al., 2008). Critically damped second order linear filters describe the time course of the population mean of post-synaptic potentials, further dispersed due to variability of parameters within each population. This mean behavior summarizes both

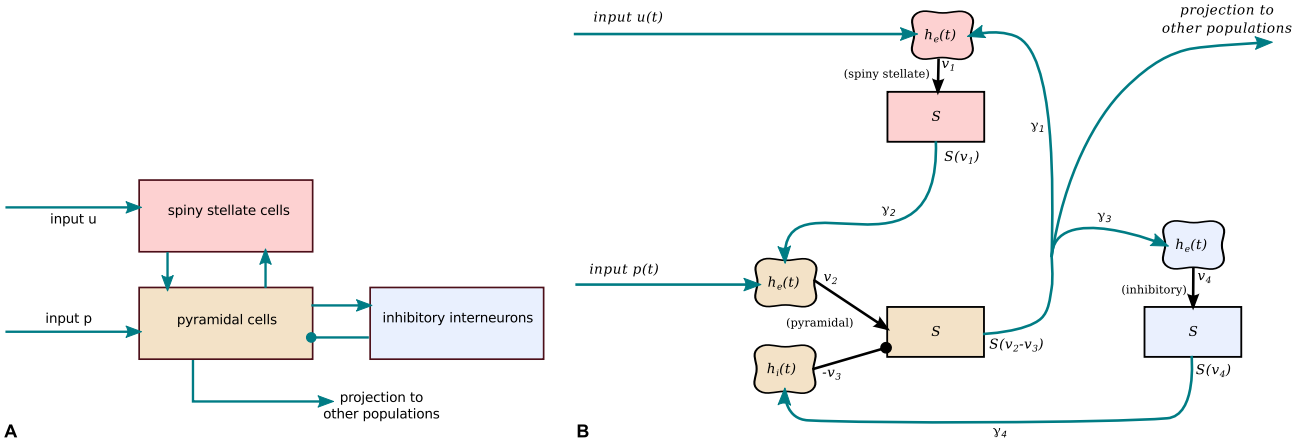


Figure 2.1: Schematic connectivity of the Jansen-Rit model. (A) Basic connectivity diagram showing the three neuronal populations, their excitatory (arrows) and inhibitory (circle) connections, and inputs from outside the local cortical region (u and p). (B) A block diagram then summarizes how this translates directly to a mathematical model: linear filter boxes labeled $h_e(t)$ and $h_i(t)$ model the mean response of excitatory and inhibitory synapse populations respectively, including their post-synaptic dendritic filtering. Sigmoid boxes (denoted S) represent conversion of mean summed soma membrane potential to mean output firing rate. Connectivity constants γ_1 to γ_4 model the number and strength of connections between populations.

synaptic and dendritic dynamics of the individual neurons. For excitatory and inhibitory synapses respectively these filters are expressed by the differential operators

$$L_e = \frac{1}{H_e \kappa_e} \left(\frac{\mathrm{d}^2}{\mathrm{d}t^2} + 2\kappa_e \frac{\mathrm{d}}{\mathrm{d}t} + \kappa_e^2 \right), \tag{2.1}$$

$$L_i = \frac{1}{H_i \kappa_i} \left(\frac{\mathrm{d}^2}{\mathrm{d}t^2} + 2\kappa_i \frac{\mathrm{d}}{\mathrm{d}t} + \kappa_i^2 \right), \tag{2.2}$$

where the scalar parameters H_e and H_i determine the maximum amplitude of the postsynaptic population response to excitatory and inhibitory inputs, respectively. The rate constants κ_e and κ_i determine the time scale of these population responses. As the synaptic filters are linear, synapses with different source or target populations can be merged where synapses are assumed to have the same aggregate properties. For example Jansen and Rit consolidated their original model to just 3 second order equations, as it was implicitly assumed that excitatory and inhibitory interneuron populations would always have identical state up to a scaling constant. Following synaptodendritic filtering, fluctuations in membrane potential sum in the cell soma and lead to changes in the average population firing rate. The sigmoid function

$$S(v) = \frac{2e_0}{1 + \exp\left[\rho_1(\rho_2 - v)\right]}, \tag{2.3}$$

describes how the mean firing rate of a neural population depends on the mean soma membrane potential v, incorporating the dispersion of responses due to variability in the parameters and underlying neuronal states (Marreiros et al., 2008). Parameters e_0, ρ_2 and ρ_1 determine the maximum firing rate, threshold potential and sensitivity, respectively.

We express the Jansen-Rit model as a set of four second-order differential equations, thus allowing both pyramidal and spiny stellate populations separately to receive extrinsic input. We follow the variable and parameter names of Moran et al. (2007). The dynamical variables v_1, v_2 and v_4 represent the positive contributions to population mean soma potentials by excitatory synapses targeting spiny stellate, pyramidal and inhibitory interneuron populations, respectively. Variable v_3 represents the negative contribution to the mean soma potential of the pyramidal population originating from inhibitory synapses. Thus the resulting mean soma potential of the pyramidal population is $v_2 - v_3$. This is taken as the main output of the model (Jansen and Rit, 1995; David et al., 2005) because the size and orientation of the apical dendrites of pyramidal neurons mean that pyramidal activity is most

closely associated with EEG signals. These equations are given by

$$L_e v_1 = \gamma_1 S(v_2 - v_3) + \langle u \rangle + \sigma_u \xi_u(t), \tag{2.4}$$

$$L_e v_2 = \gamma_2 S(v_1) + \langle p \rangle + \sigma_p \xi_p(t), \tag{2.5}$$

$$L_i v_3 = \gamma_4 S(v_4), \tag{2.6}$$

$$L_e v_4 = \gamma_3 S(v_2 - v_3). \tag{2.7}$$

Equation (2.4) describes excitatory synaptic input targeting the spiny stellate population. Equations (2.5) and (2.6) describe excitatory and inhibitory synaptic input to the pyramidal population, respectively. Equation (2.7) describes excitatory synaptic input to the inhibitory interneuron population. Parameters $\langle u \rangle$ and $\langle p \rangle$ are the mean per-neuron external input firing rates to the cortical region, targeting spiny stellate and pyramidal populations, respectively. Langevin white noise terms $\xi_u(t)$ and $\xi_p(t)$ in the extrinsic input represent the fluctuations in the input firing rates, with σ_u and σ_p denoting their standard deviations. Scalar connectivity constants γ_1 to γ_4 represent at the population scale the number and strength of connections between the three neural populations.

This system of equations is equivalent to a single 8-dimensional stochastic first order differential system:

$$\mathbf{dv} = \mathbf{f}(\mathbf{v})\, dt + G\, \mathbf{dW}(t) \tag{2.8}$$

where matrix elements of G determine the cross correlation of noise inputs to the pyramidal and spiny stellate populations. This is the equation that we integrate numerically.

Table 2.1 lists the values of parameters used for all simulations in this study; they are the standard parameter values introduced by Jansen and Rit (1995). Jansen and Rit themselves focused on numerical simulations of this nonlinear model. Through a survey of the simulated behavior with physiologically realistic parameters, they observed a variety of noise-driven rhythmic behaviors consistent with human alpha and beta rhythms. Wendling et al. (2000) studied the emergence of "spike-wave" oscillations resembling epileptic activity when the ratio of excitation to inhibition was increased. Bifurcations in this model where subsequently examined by Grimbert and Faugeras (2006) who treated the input p as the bifurcation parameter in order to understand better the original simulation results of Jansen and

Table 2.1: Jansen-Rit standard parameter values

Parameter	Value	Description
H_e	$3.25\,\text{mV}$	maximum amplitude of the excitatory postsynaptic population response
H_i	$22.0\,\text{mV}$	maximum amplitude of the inhibitory postsynaptic population response
κ_e	$100\,\text{s}^{-1}$	rate constant for postsynaptic population response to excitatory input
κ_i	$50\,\text{s}^{-1}$	rate constant for postsynaptic population response to inhibitory input
e_0	$2.5\,\text{s}^{-1}$	half of the maximum population mean firing rate
ρ_2	$6.0\,\text{mV}$	population mean firing threshold potential
ρ_1	$0.56\,\text{mV}^{-1}$	firing rate sigmoid function voltage sensitivity parameter
γ_1	135	connectivity constant: pyramidal to spiny stellate
γ_2	108	connectivity constant: spiny stellate to pyramidal
γ_3	33.75	connectivity constant: pyramidal to inhibitory interneurons
γ_4	33.75	connectivity constant: inhibitory interneurons to pyramidal

Rit (limit cycle beyond a Hopf bifurcation causing alpha oscillations) and Wendling et al. (the emergence of a large amplitude non-harmonic oscillator near a sniper bifurcation). More recently, Spiegler et al. (2010) performed a more general bifurcation analysis that included time scale parameters and analyzed the presence of qualitatively different oscillatory regimes.

2.2.2 Bifurcation parameters

In the original model (Jansen et al., 1993), both pyramidal and excitatory interneuron populations were the targets of extrinsic inputs, with the two inputs being always proportional (fully correlated). In the model of Jansen and Rit (1995), all extrinsic input was delivered to the pyramidal neurons only, with external stimulation of the other population dropped.

David and Friston (2003), revisited the Jansen-Rit model, in particular explicitly identifying the "excitatory interneuron" population of the original model with spiny stellate cells in layer IV of the neocortex. Their motivation was to send extrinsic input to the layer IV spiny stellate cells in the model

rather than to the pyramidal cells. This was arguably a more realistic model of connectivity for input representing thalamocortical sensory afferents. However the equations as published retained the pyramidal-only input of the original Jansen-Rit model.

Moran et al. (2007), in the context of Dynamic Causal Modeling (DCM, a framework for model selection and parameter estimation), extended the Jansen-Rit model with several innovations, including firing rate adaptation, recurrent inhibition, and a differently shaped sigmoid function. In particular Moran et al. did change the target of the extrinsic input to be the spiny stellate population, as foreshadowed by David and Friston. We refer to this model as the Moran-Friston model hereafter.

For the present study we minimally extend the Jansen-Rit model, so that extrinsic input can be delivered either to the pyramidal population (as in Jansen and Rit, 1995), the spiny stellate excitatory population (as in Moran et al., 2007) or more realistically a combination of the two. In this way, the system input can be varied continuously from the Jansen-Rit design to the Moran-Friston design or anywhere in between. In addition, for the case of input to both populations, these two inputs can be chosen as uncorrelated, fully correlated or partially correlated in their fluctuations. Hence we study the bifurcations of this model as input is varied in the combined (u, p) plane. This subset of parameter space includes a one-dimensional space explored by Jansen and Rit containing a supercritical Hopf bifurcation studied by Grimbert and Faugeras that is within the physiological range of parameters. We present the bifurcation analysis in Sec. 2.3.1.

2.2.3 Numerical simulation and analysis

The model is a system of stochastic differential equations (SDEs) with additive noise. Equation (2.8) is integrated numerically using the Heun algorithm, which is applicable to SDEs in Stratonovich form (Rümelin, 1982). This ensures that noise amplitude is scaled in appropriate proportion to the square root of the integration time step. We use an integration time step of 0.2 ms. The first transient 5 seconds of each simulation is discarded from further analysis.

As reviewed in the introduction, the approach to linear instabilities in systems of equations such as (2.4–2.7) is widely assumed to cause changes in the autocorrelation length and/or a peak in the power spectral density function (in the case of a Hopf bifurcation). This is because it is often assumed that

the linear treatment of these systems — which predicts both an enhancement of spectral peaks and a lengthening of the autocorrelation time — can be extrapolated from the setting when the system is linearly stable to when it is in the neighborhood of a bifurcation.

To estimate the normalized autocorrelation function of the resulting time series we first normalize each time series to a mean of 0 and standard deviation of 1, and then compute the cross-correlation of the series with itself applying unbiased normalization,

$$\hat{R}_{xx,\text{unbiased}}(m) = \frac{1}{N-m} \sum_{n=0}^{N-m-1} x_{n+m} x_n \quad (m \geq 0), \tag{2.9}$$

where m is the lag expressed as number of samples (Orfanidis, 1996). In each case we compute autocorrelation at lag times from 0 to one quarter of the total time series length for further analysis. Since time series are generated in the vicinity of Hopf bifurcations with natural frequency about 11 Hz, the autocorrelation functions all have a strong 11 Hz component. Because we are primarily interested in the decay of the amplitude of this autocorrelation over a longer time scale, each simulation is repeated 16 times with identical parameters to generate 16 sample paths each of duration 600 s or 1800 s. The autocorrelation function is calculated as described above for each sample path separately. The decay is then quantified by calculating the modulus of the Hilbert transform of the normalized autocorrelation functions computed above. The pointwise mean and standard deviation of this autocorrelation amplitude across 16 sample paths are then plotted. Power spectra are estimated using the Welch algorithm with Hamming window and a segment length of 80 000 samples or 16 seconds.

The full MATLAB code implementing the model, integration and time series analyses is available from the authors on request.

2.3 Results

2.3.1 Bifurcation diagram

From the earlier bifurcation analysis of Grimbert and Faugeras (2006) the model is known to have a supercritical Hopf bifurcation when the pyramidal input $p = 89.8\,s^{-1}$ and the other parameters

are set to the values used by Jansen and Rit (1995). This assumed no input to the excitatory (spiny stellate) population. We label this bifurcation point H1; it has mean input firing rate $\langle p \rangle = 89.8\, s^{-1}$ to the pyramidal population and zero input to the spiny stellate population (i.e., $\langle u \rangle = 0, \sigma_u = 0$). This maps directly back to the original Jansen-Rit model with pyramidal-only input. Matching the effective noise level used by Jansen and Rit (correcting a scaling error in the original paper) can be achieved by allowing p to fluctuate with standard deviation approximately $\sigma_p = 2.0\, s^{-1}$. Our simulations below use a value of $\sigma_p = 0.5390\, s^{-1}$ or roughly one quarter of the noise level of Jansen and Rit.

To examine the difference between cases where input is provided in different ratios to the spiny stellate population and pyramidal population, we continue the bifurcation point H1 in the (u, p) plane in parameter space, using the numerical continuation package MATCONT (Dhooge et al., 2003). Figure 2.2 shows the bifurcation diagram in the (u, p) plane. This plane is a 2-dimensional slice through the larger parameter space of the model, so that a curve in this plane corresponds to a surface in parameter

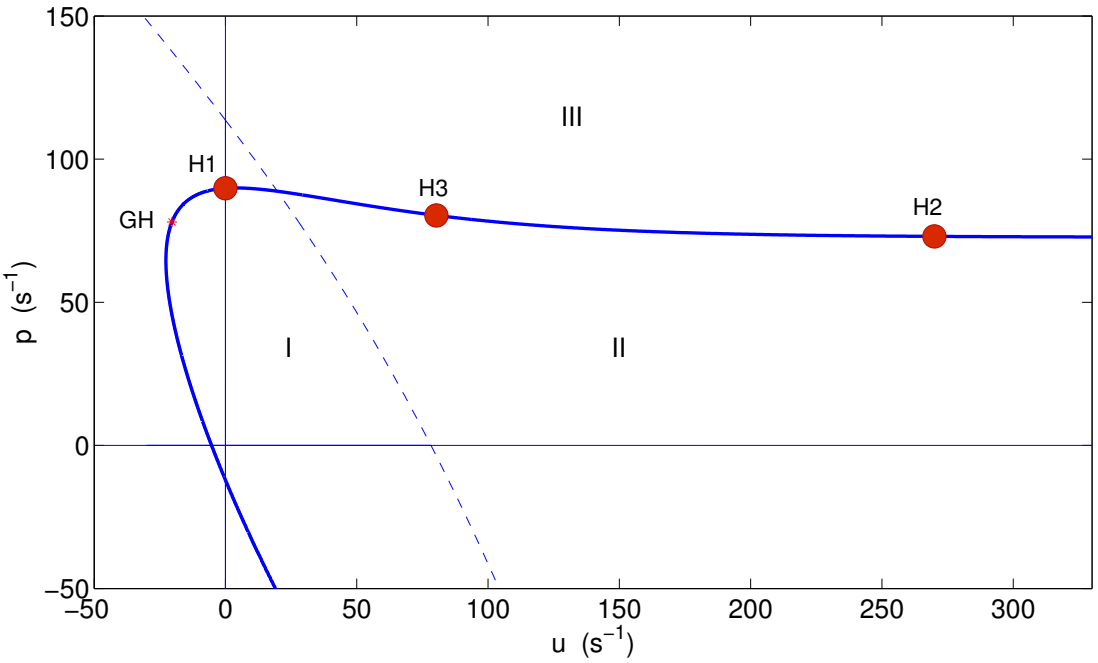


Figure 2.2: Bifurcation diagram in the (u, p) parameter plane showing Hopf curve (thick, solid curve) and the location of the chosen Hopf bifurcation points H1, H2 and H3 on this curve. A Generalized Hopf bifurcation (GH) marks the transition from subcritical Hopf points (the curve below GH) to supercritical Hopf points (the curve continuing beyond GH). Regions where $p < 0$ or $u < 0$ are non-physical. Below the Hopf curve (regions I and II) a stable fixed point exists, which gradually loses linear stability as the curve is approached. Above the Hopf curve (region III) this point has lost linear stability and become a stable limit cycle. The dashed line is a curve of fold bifurcation points. In region II a single stable fixed point exists. In region I the system is bistable with a second stable fixed point also existing, at lower excitation.

space. The Hopf curve is almost horizontal for $u > 0$, implying that the level of pyramidal cell stimulation required to reach the supercritical Hopf bifurcation in the model ($p \sim 75$–$90 s^{-1}$) is roughly independent of the level of spiny stellate cell stimulation for $u \gtrsim 0$. For comparison with H1, we select point H2 on this same surface of supercritical Hopf points, but this time with greater mean input to the spiny stellate population ($\langle u \rangle = 270\, s^{-1}$) than to the pyramidal population ($\langle p \rangle = 73\, s^{-1}$). The magnitude of fluctuations $\sqrt{\sigma_u^2 + \sigma_p^2}$ in the input is kept the same as at H1, with a standard deviation of $\sigma_u = 0.5203\, s^{-1}$ in the spiny stellate input and $\sigma_p = 0.1407\, s^{-1}$ in the pyramidal input.

2.3.2 Autocorrelation indicator behaves differently at H1 and H2

For each of the bifurcation points we simulate the dynamics at four locations in parameter space: the approach to the bifurcation from the linearly stable side (two points), at the bifurcation point (one point), and beyond the bifurcation (one point). In each case the output pyramidal time series ($v_2 - v_3$) is the focus of our analysis.

Each simulation is performed separately with parameter values fixed at these different values, rather than performing a single dynamic simulation with sliding parameters. This approach allows the time series analyzed at a fixed parameter value to be approximately stationary (provided the total time simulated is long enough) so that statistics for the process at that parameter point can be estimated from a finite time series. Where variance is considered as an indicator of instability this approach also avoids any spurious short-time increases in variance due to the dynamically shifting range of the system in phase space, as distinct from increased noise-driven variance at the new parameter values (Kuehn, 2011).

To determine the effect of proximity to a bifurcation on the fluctuation statistics, we analyze the approach and passage through bifurcations H1 and H2. Figure 2.3 shows the results for bifurcation H1. An exemplar pyramidal time series [Figure 2.3(A)] reveals a fluctuating oscillatory system, whose power spectrum [Figure 2.3(B)] peaks at the frequency of the Hopf instability, namely 11 Hz. The series of panels in Figure 2.3(C) shows that when approaching and passing point H1 (from left-to-right), the autocorrelation time stays approximately constant.

For comparison, the corresponding analyses for bifurcation H2 are shown in Figure 2.4. By eye, the fluctuation envelope of the amplitude appears smoother. As is evident in Figure 2.4(C), the autocorrelation amplitude decays much more slowly as the system approaches the bifurcation. At the bifurcation point H2 the autocorrelation remains above 20% of its zero-lag value at a lag of 15 s. The variance of the output pyramidal time series increases as the bifurcation H1 is approached, with standard deviations of 0.4550 mV, 0.5344 mV, 0.5630 mV and 0.6110 mV at the four parameter points respectively. Approaching H2 this also occurs, with standard deviations of 0.1454 mV, 0.2160 mV, 0.2582 mV and 0.3449 mV respectively for the output time series. It is notable that in the vicinity of point H2, the standard deviation of the simulated pyramidal output time series is between 1.8 and 3.1 times smaller than in the vicinity of point H1, while autocorrelation times are roughly 7 times longer than at H1.

As expected for a Hopf bifurcation in a stochastic system the dynamics change gradually and continuously through the bifurcation (Rowat and Greenwood, 2011). The amplitude of oscillations increases when moving towards and beyond the bifurcation point as revealed by increased variance of the output time series. Close to the bifurcation point this reflects weakening of the stability of the (fixed point) attractor while beyond the bifurcation point it reflects increasing size of the (limit-cycle) attractor. The increase in power at 11 Hz is visible in the power spectrum (Figure 2.3(B)).

The comparison of Figures 2.3(C) and 2.4(C) shows that autocorrelation is a useful and clearly visible indicator of linear instability in the vicinity of point H2, but not for point H1. This is despite these being points on the same surface of bifurcations with the same variance of input fluctuations.

We conjecture that the key difference between H1 and H2 is the orientation of the input fluctuations in phase space with respect to the two-dimensional center manifold of the bifurcation, which determines the specific directions in which linear stability is weakening. When close to the equilibrium point, the center manifold surface can be approximated by the center eigenspace of the bifurcation. Since the eigenvectors of the linearized system are far from orthogonal the relevant reference plane to determine the noise component projected into the center eigenspace is that plane perpendicular to the stable eigenspace. For H1 the resulting projection of the noise vector onto this reference plane is $\cos\alpha = 0.0031$. For H2 the projection is $\cos\alpha = 0.0008$; i.e., the noise input has a projection onto that plane that is four times larger in the case of H1 than in the case of H2.

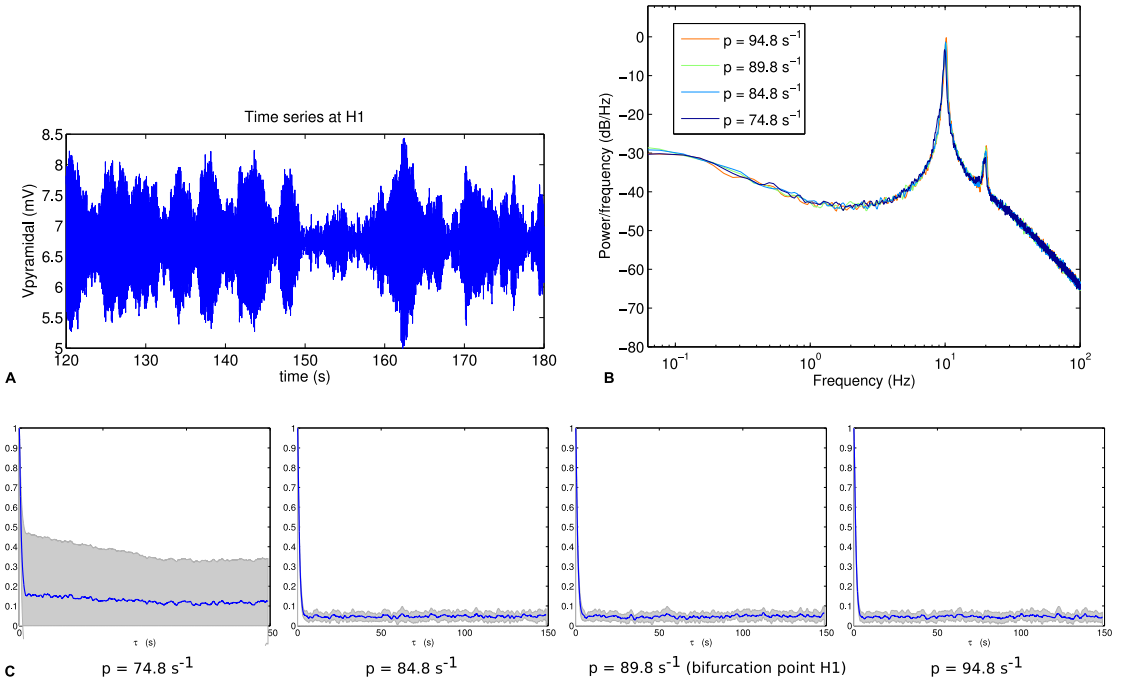


Figure 2.3: (A) Example section of simulated time series at the bifurcation point H1. (B) Power spectrum at H1. (C) Autocorrelation amplitude at points in parameter space approaching the bifurcation point H1 ($p = 74.8\,s^{-1}$, $p = 84.8\,s^{-1}$), at the bifurcation point H1 ($p = 89.8\,s^{-1}$) and beyond the bifurcation point ($p = 94.8\,s^{-1}$). The line indicates the mean over 16 trials and the grey area indicates one standard deviation.

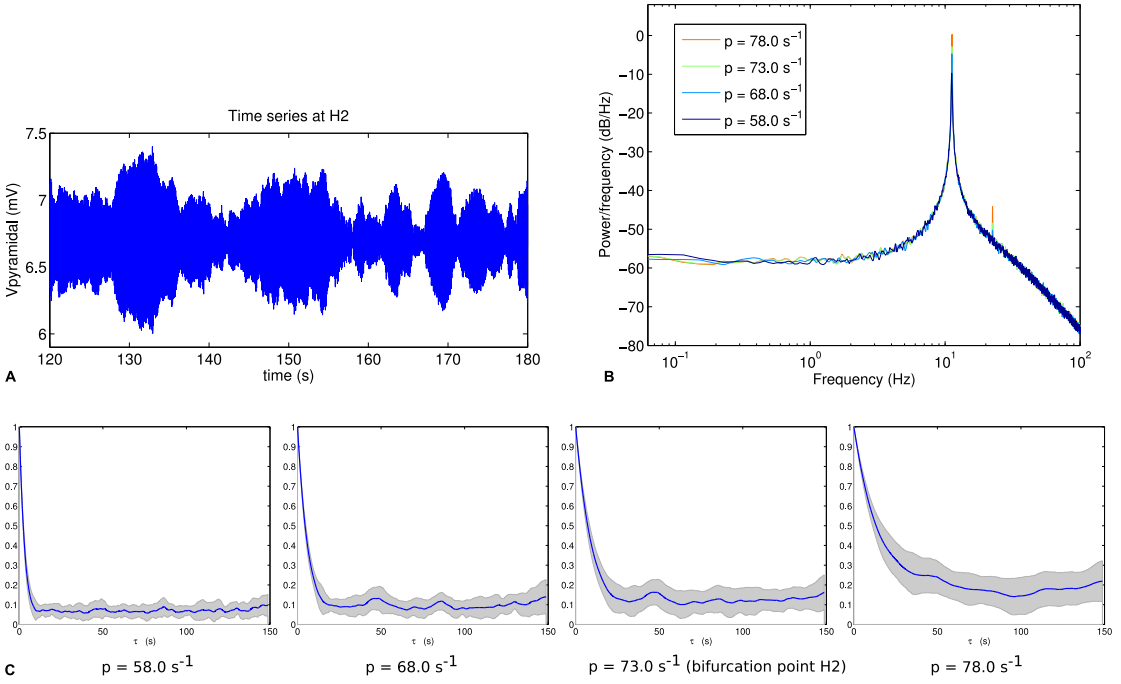


Figure 2.4: (A) Example section of simulated time series at the bifurcation point H2. (B) Power spectrum at H2. (C) Autocorrelation amplitude at points in parameter space approaching the bifurcation point H2 ($p = 58.0\,s^{-1}$, $p = 68.0\,s^{-1}$), at the bifurcation point H2 ($p = 73.0\,s^{-1}$) and beyond the bifurcation point ($p = 78.0\,s^{-1}$). The line indicates the mean over 16 trials and the grey area indicates one standard deviation.

However, the comparison between points H1 and H2 does not by itself give strong support for this hypothesis, because there are several other factors that are significantly different between H1 and H2. In particular H2 has 3.1 times the total input firing rate of H1, so that on this basis the difference in autocorrelation could simply be due to greater level of excitation for point H2. This motivates the comparison constructed below.

2.3.3 Autocorrelation depends on orientation of input fluctuations

In order to separate the effect of different mean firing rates from the effect of different noise orientation, we construct two new scenarios H3p and H3u, where the only difference between them is the noise orientation; all other parameters are kept identical. We choose point H3 on the same bifurcation line of supercritical Hopf points, but with equal mean input firing rates to pyramidal and spiny stellate populations (mean input firing rate per neuron of $\langle u \rangle = \langle p \rangle = 80.35\, s^{-1}$). We simulate two scenarios at point H3 to test the conjecture, with both scenarios using the same values for all model parameters, and in particular with both scenarios using the same mean input firing rates, as illustrated in Figure 2.5. We define the scenario H3p as the case where only the pyramidal input is allowed to fluctuate about

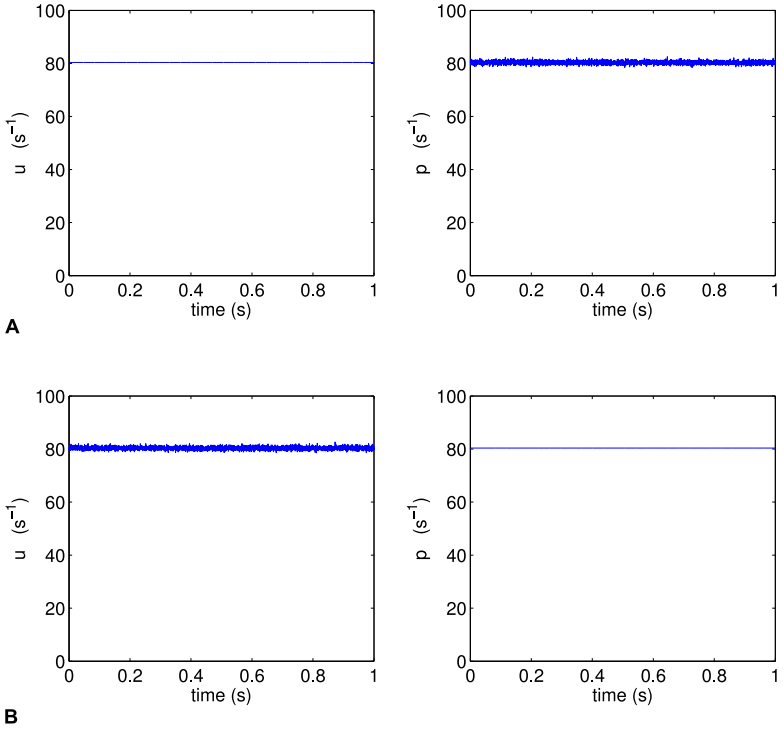


Figure 2.5: One second of sample external inputs u and p, (A) for scenario H3p and (B) for scenario H3u

its mean, while spiny stellate input is held steady at its mean value, corresponding to the parameters $\langle p \rangle = 80.35\, s^{-1}$, $\sigma_p = 0.5390\, s^{-1}$, $\langle u \rangle = 80.35\, s^{-1}$ and $\sigma_u = 0\, s^{-1}$.

Scenario H3u is defined as the case where only the spiny stellate input is allowed to fluctuate, while pyramidal input is held steady, corresponding to the parameters $\langle p \rangle = 80.35\, s^{-1}$, $\sigma_p = 0\, s^{-1}$, $\langle u \rangle = 80.35\, s^{-1}$ and $\sigma_u = 0.5390\, s^{-1}$.

By using these two constructed scenarios, all parameters in the simulation are kept identical between scenarios H3p and H3u except for the direction of the fluctuations of input in phase space, which is rotated in phase space from the pyramidal direction to the spiny stellate direction. Rotating the vector of fluctuations independently from the vector of mean inputs is non-physiological. The simulated results of the non-physiological scenarios H3p and H3u are used to shed light on the reason for different autocorrelation in the original realistic scenarios H1 and H2.

Comparable analyses of these two scenarios are presented in Figures 2.6 and 2.7. The contrast between scenarios H3p and H3u is clear in Figures 2.6(C) and 2.7(C). When the fluctuations are in the input to the pyramidal population (scenario H3p), the decay of autocorrelation amplitude changes little as the bifurcation point is approached. By contrast, when the fluctuations are in the input to the spiny stellate population (scenario H3u) the indicator of increased autocorrelation length is very prominent. A large increase in autocorrelation heralds the transition to linear instability in scenario H3u with significant autocorrelation at lags of up to 450 s. This indicator is much less evident in scenario H3p.

It is also instructive to view the autocorrelation amplitude with log scaling of the delay axes. The results for the four scenarios (H1, H2, H3p, H3u) we have thus far considered are shown in Figure 2.8. Whereas the autocorrelation length stays almost invariant across the bifurcation in scenario H1 (panel A), a clear increase is seen in scenario H2 (panel B). Where scenario H3p shows a small, but systematic lengthening (panel C), a progression through the same points in parameter space — but now with input fluctuations aligned with the stable eigenspace — can again be seen to lead to a dramatic increase (panel D).

In both scenarios H3p and H3u the variance of the output pyramidal time series increases as the bifurcation point H3 is approached, with standard deviations of 0.4884 mV, 0.4730 mV, 0.5134 mV and 0.5339 mV for H3p and 0.0188 mV, 0.0295 mV, 0.0873 mV and 0.3033 mV for H3u. The standard deviation is starkly different between these two scenarios, with standard deviation between 2 and 22

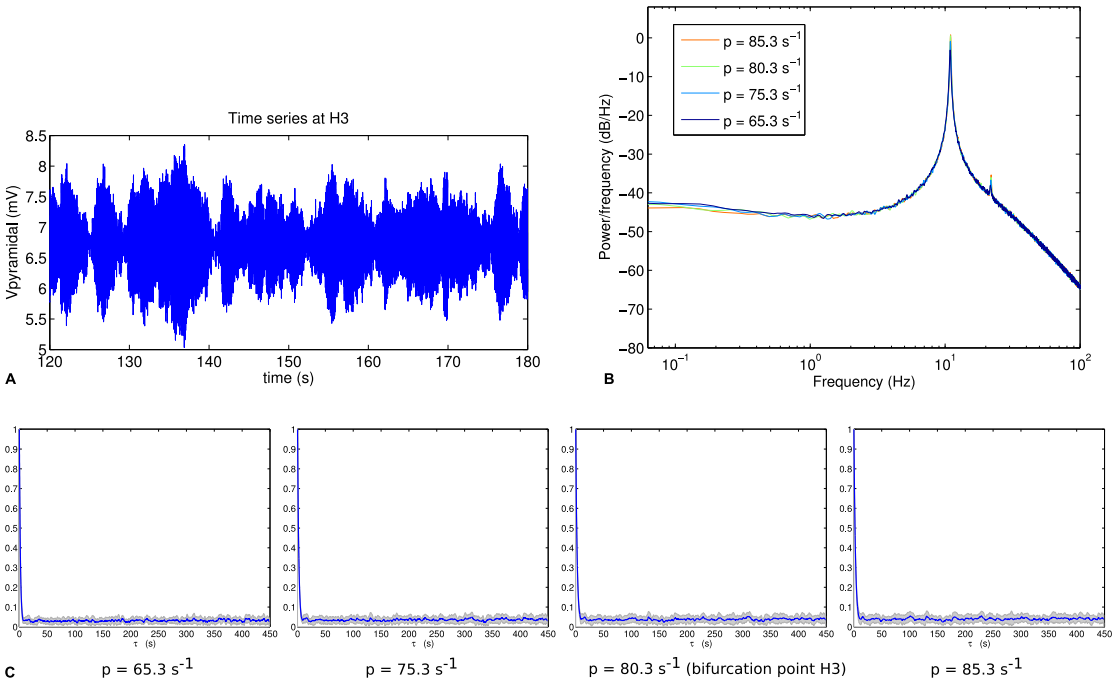


Figure 2.6: (A) Example section of simulated time series for scenario H3p. (B) Power spectrum at H3. (C) Autocorrelation amplitude at points in parameter space approaching the bifurcation point H3 ($p = 65.3\,s^{-1}$, $p = 75.3\,s^{-1}$), at the bifurcation point H3 ($p = 80.3\,s^{-1}$) and beyond the bifurcation point ($p = 85.3\,s^{-1}$). The line indicates the mean over 16 trials and the grey area indicates one standard deviation.

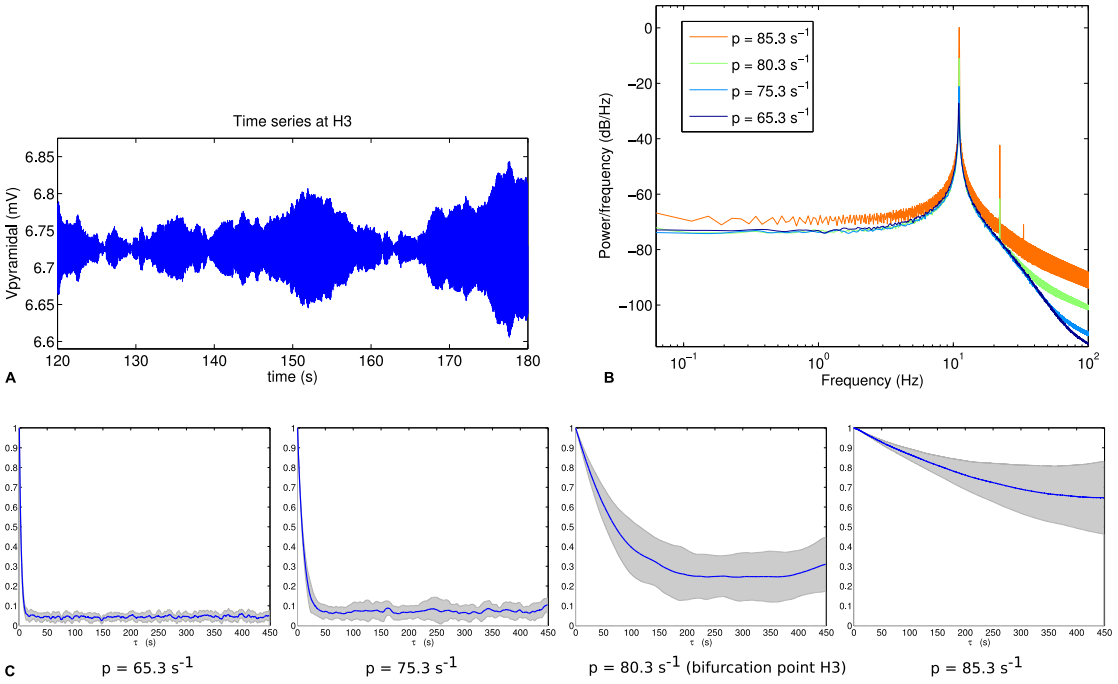


Figure 2.7: (A) Example section of simulated time series for scenario H3u. Note the y-axis scale is much smaller than that of Figure 6(A), reflecting much smaller output variance in this case. (B) Power spectrum at H3. (C) Autocorrelation amplitude at points in parameter space approaching the bifurcation point H3 ($p = 65.3\,s^{-1}$, $p = 75.3\,s^{-1}$), at the bifurcation point H3 ($p = 80.3\,s^{-1}$) and beyond the bifurcation point ($p = 85.3\,s^{-1}$). The line indicates the mean over 16 trials and the grey area indicates one standard deviation.

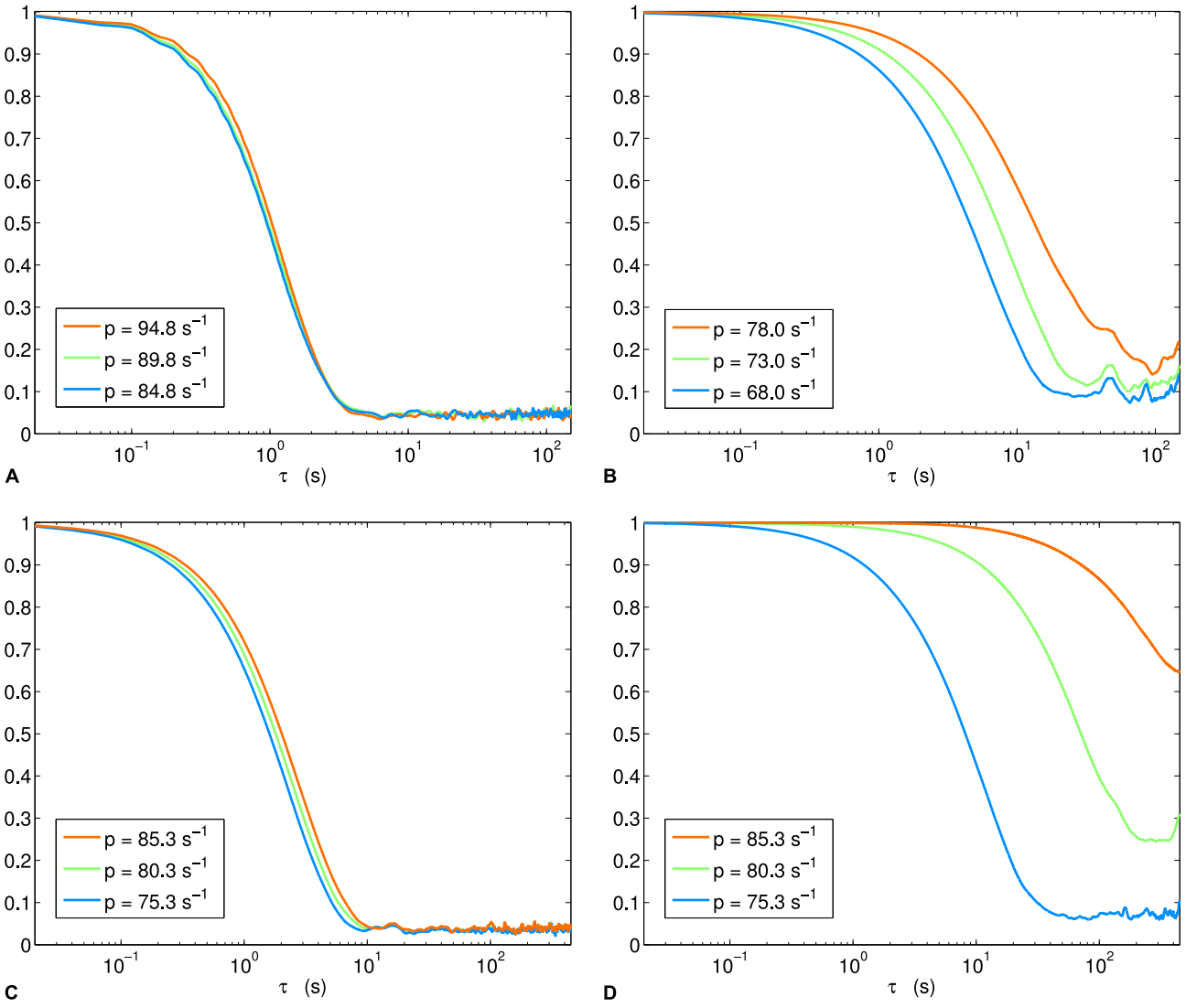


Figure 2.8: Autocorrelation amplitude in log(delay)-linear(correlation) coordinates. Each panel shows step immediately before (blue), at (green) and beyond (orange) Hopf bifurcation. (A) Scenario H1, (B) H2, (C) H3p, (D) H3u.

times smaller in scenario H3u than in scenario H3p. Thus changing the noise input direction results in both reduced variance and increased autocorrelation. Variance also increases more rapidly in scenario H3u than H3p as the bifurcation is approached.

Relating this to the orientation of input noise, the contrast in alignment is even greater between scenarios H3p and H3u than in the comparison of H1 and H2. For H3p the projection of the input noise onto the reference plane perpendicular to the stable eigenspace is $\cos\alpha = 0.003$, i.e., noise input has a non-negligible component perpendicular to the stable eigenspace of the bifurcation near the equilibrium point, whereas for H3u the projection is $\cos\alpha = 0.00006$, i.e., the projection of the noise input onto the reference plane is fifty times smaller in the case of H3u than H3p.

2.3.4 Input correlation and output variable not important

The results presented above are calculated from the pyramidal time series. Applying the same process to time series for the other two populations in the model (spiny stellate and inhibitory) shows that in each case, the results for autocorrelation decay and variance show the same behavior as the pyramidal time series. This is important, as it rules out the possibility that the autocorrelation difference results from different amounts of filtering between the noise input and the measured output.

In the case of point H2, both pyramidal input and spiny stellate input have fluctuations. To check whether correlations between the two input fluctuations are important to the results we examine the two extreme cases of independent and perfectly correlated inputs. Cross-correlation of input fluctuations does not affect the results: autocorrelation amplitude of the pyramidal output decays over a similar time scale whether the inputs to the two populations are independent or perfectly correlated.

2.3.5 Scaling properties of output fluctuations

As reviewed earlier, long-tailed fluctuation distributions have been observed in the amplitude fluctuations of alpha (Freyer et al., 2009) and beta oscillations in scalp EEG data (Linkenkaer-Hansen et al., 2004). Therefore we study the statistical properties of fluctuations at and near the bifurcation points in scenarios H3u and H3p. In particular, we characterize fluctuations by the distributions of sizes and durations of excursions in the amplitude envelope of the detrended pyramidal time series. More specifically, we analyze the squared Hilbert amplitude, which is a measure of instantaneous power. We extract excursions above a threshold (sometimes termed “avalanches” in the literature) where each excursion is delineated by the time points at which the instantaneous power crosses the threshold from below and the next crossing from above. Fluctuation duration is thus the length of the time interval for which the power is above threshold, and we define fluctuation size to be the time integral of the instantaneous power over this interval (i.e., the area under the curve, a measure of energy in the fluctuation). We choose the threshold for each time series such that it approximately maximizes the number of identified fluctuations and falls in a regime where the fluctuation statistics are relatively insensitive to small changes in this value.

We analyze the fluctuation size and duration distributions following the methods of Clauset et al. (2009). For each set of fluctuation statistics we calculate the inverse cumulative distribution function and fit candidate distributions to the tail using the method of maximum likelihood: power law (the Pareto distribution), power law with exponential cutoff, lognormal, and exponential. Here the tail is all the data above a lower bound that we identify as the value that minimizes the Kolmogorov-Smirnov goodness-of-fit statistic between the power law model and the data (Clauset et al., 2009). This method of determining the range of the fit from the data strikes a balance between fitting too wide a range (i.e., outside the power-law regime) and too narrow a range (i.e., throwing away data unnecessarily). We use the same fitting range for all four candidate distributions. We estimate a p-value for the fitted power law by comparing the data to 1000 synthetic data sets drawn from a true power law, which accounts for whether the deviation between the data and the fitted power law is within the range expected for finite sampling of a true power law. The p-value is taken to be the fraction of synthetic data sets that deviate from the power law by at least as much as the data, and $p > 0.1$ indicates plausibility of the power law hypothesis (Clauset et al., 2009). We compare the fitted power law with alternative distributions using likelihood ratio tests. Significant deviation of the likelihood ratio from zero is tested using Vuong's methods (Vuong, 1989). For the nested hypothesis of power law versus power law with cutoff (the latter family includes the former), the null hypothesis is that the power law is best-fitting distribution. For all other tests, the null hypothesis is that both distributions are equally far from the true distribution.

Figure 2.9 shows the fluctuation distributions for H3u. The empirical distributions for both duration (Figure 2.9A) and area (Figure 2.9B) exhibit a scaling regime over approximately four orders of magnitude. The power law fits for duration and area have exponents 1.56 and 1.51, and p-values $p = 0.27$ and $p = 0.72$, respectively, and are thus consistent with the hypothesis that the true distribution is a power law. The fitted exponents depend weakly on the threshold value but the main finding of a broad scaling regime is unchanged. The lognormal and power law with exponential cutoff are also consistent with the data: the likelihood ratio tests do not distinguish between the lognormal and power law fits (duration: $p = 0.15$; area: $p = 0.26$), but favor the power law with exponential cutoff over both power law (duration: $p = 0.016$; area: $p = 0.047$) and lognormal (duration: $p = 0.003$; area: $p = 0.004$). The pure exponential distribution is strongly ruled out in all cases ($p \ll 0.001$) and so is not shown.

Approach to this bifurcation, shown in Figure 2.10, reiterates the autocorrelation results of Section

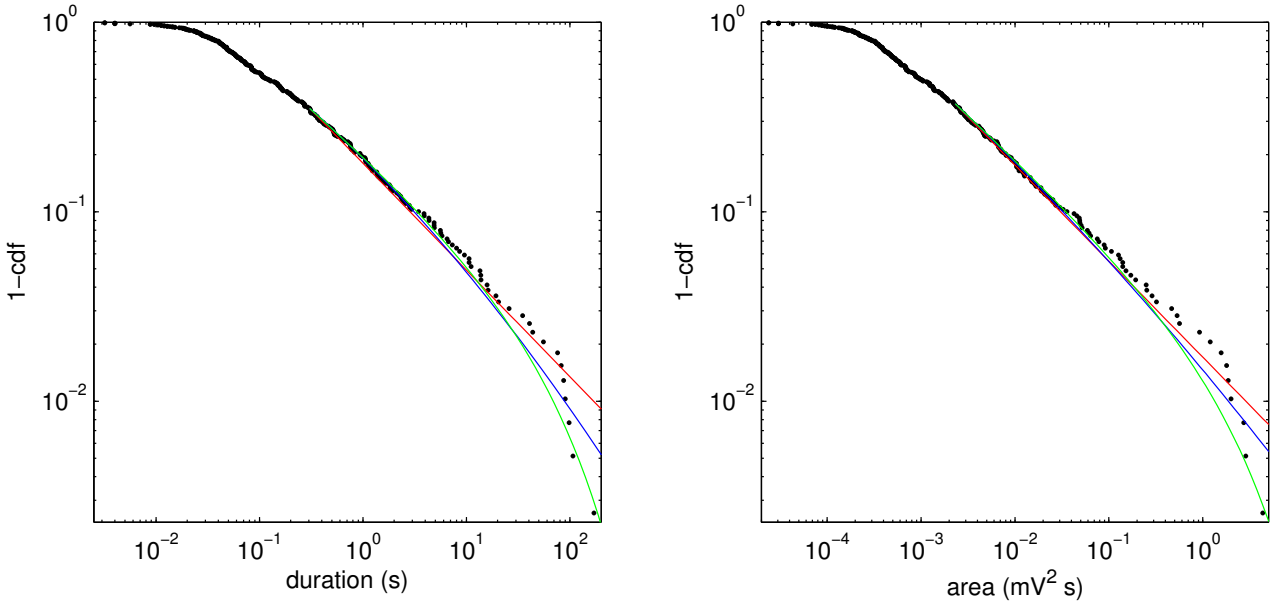


Figure 2.9: Upper cumulative distributions of fluctuation statistics at the bifurcation point H3u, using squared Hilbert amplitude thresholded at $0.008 mV^2$, with power law (red), power law with exponential cutoff (green), and lognormal (blue) fits plotted for the fitted range of the tail. (A) Fluctuation duration. (B) Fluctuation size as given by area under the curve.

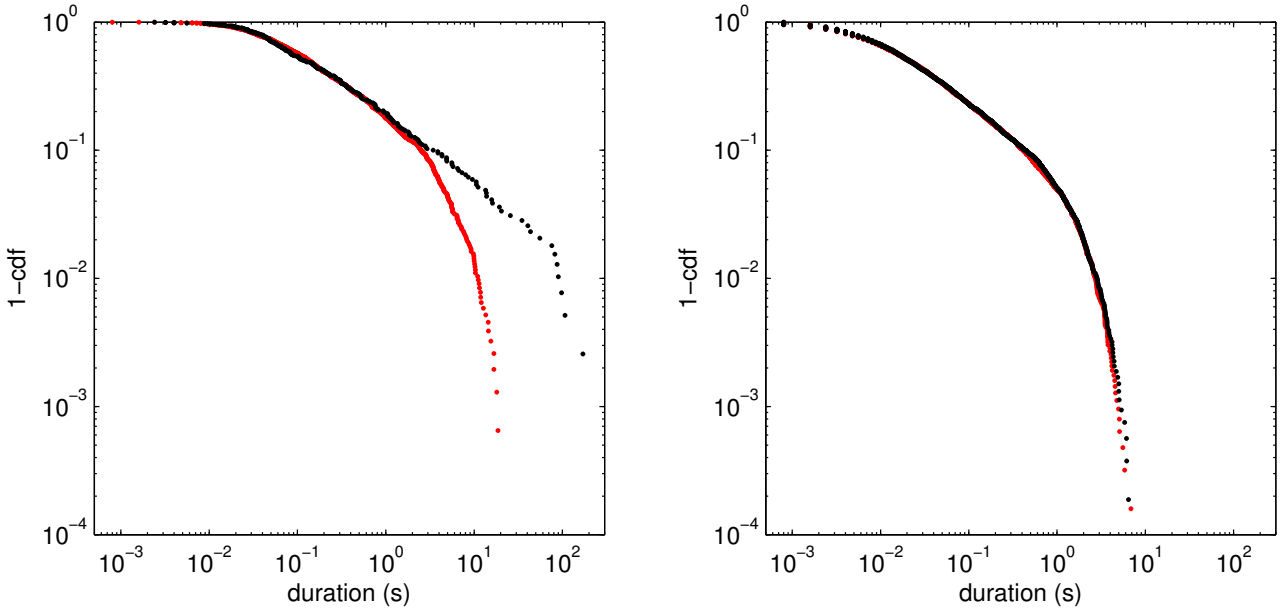


Figure 2.10: Comparison of fluctuation duration distributions between points approaching ($p = 65.3\, s^{-1}$) and at the bifurcation point ($p = 80.3\, s^{-1}$) for the two noise input directions. (A) Point H3u (black, threshold=$0.008\, mV^2$) and nearby more stable point (red, threshold=$0.0002\, mV^2$). (B) Point H3p (black, threshold=$0.4\, mV^2$) and nearby more stable point (red, threshold=$0.18\, mV^2$).

2.3.3. Near H3u (Figure 2.10A), the long scaling regime of Figure 2.9A (black) is significantly diminished away from the bifurcation (red), with few fluctuations having durations > 10 s. Here, the pure power law is ruled out ($p < 0.001$), and the power law with exponential cutoff is strongly favored over all the alternatives tested. For comparison, Figure 2.9B shows fluctuations at the same bifurcation when noise enters almost perpendicular to the center eigenspace (scenario H3p). At the bifurcation (black), there is no clear scaling regime, and the distribution is essentially unchanged by moving to a more stable point in parameter space (red). The power law fit is ruled out for both points ($p < 0.001$), and again the power law with cutoff is strongly favored. Thus, as in the autocorrelation cases (Figures 2.7C and 2.8C), the fluctuation statistics clearly herald the approach to the bifurcation for H3u but only negligibly for H3p.

2.4 Discussion

The relevance of these results to physiology is twofold. Firstly, we have demonstrated a fundamental limitation in the use of autocorrelation as an indicator of the loss of linear stability, a limitation which will apply when attempting to detect bifurcations from actual human EEG, EMG and MEG time series. Secondly, the demonstration of both long autocorrelation times and scale-free temporal fluctuations in a simple, low dimensional stochastic model informs the debate about whether the brain exhibits self-organised criticality, because it shows that these features can also arise from mechanisms other than a multi-scale critical phase transition.

Close to the supercritical Hopf bifurcation in the Jansen-Rit model, we have shown that when lengthened autocorrelation times and scale free fluctuations manifest in any one cell population as indicators of approach to the bifurcation, then they are indeed detectable in the pyramidal time series that is most closely associated with EEG signals. The standard parameters of the model provide sufficiently large coupling between the three neural populations that lengthened autocorrelation is evident in all three populations when it is present in any of them.

When considering long time autocorrelation and scale free fluctuations that are present in human EEG time series this suggests that in addition to the possibility that these could arise in the brain at the point of phase transition in a complex, multi-scale system (Stam and de Bruin, 2004; Linkenkaer-Hansen

et al., 2004), there may also be a role for low-dimensional stochastic dynamics at the population scale in generating these indicators.

More importantly, we have shown that even in a very simplistic cortical model, these indicators can already be subtle in their dependency on neuronal inputs. Longer autocorrelation times are not guaranteed to be evident in the output just because there is a bifurcation where linear stability is lost. In particular we have shown that a change of the orientation in phase space of small fluctuations in the input can be sufficient to enhance or almost completely remove this indicator.

Jansen and Rit suggested that input to excitatory interneurons could be removed from the model, as input to the pyramidal population from coupled columns was expected to have the same effect (Jansen and Rit, 1995). Our results show that when fluctuations in the input are taken into account, the statistical properties of the model output are sensitive to the choice of which neural population receives the extrinsic input.

The Jansen-Rit model is representative of a broad class of models that mathematically can be expressed as a composition of sigmoid functions and second order linear filters. It is worth noting that neural field models (such as Jirsa and Haken (1996) and Robinson et al. (1997)), when restricted to spatially uniform solutions, can also be expressed in this mathematical form, with an additional critically damped second order linear filter capturing the time characteristics of local axonal propagation with a population spread of sources and axon parameters (Robinson et al., 1997). There are no particularities of the current model that suggest that the phenomena which we describe will be limited to this setting. The present results regarding fluctuation orientation hence speak broadly to the commonly employed neural mass and neural field models of large-scale neuronal activity.

2.4.1 Opposite effect on autocorrelation and variance

Autocorrelation and variance of the output signal have been suggested as generic indicators of the approach to local bifurcation, as standard linear analysis shows they are both expected to increase as the bifurcation is approached and the real part of bifurcating eigenvalues approaches zero. We also observed that changing the orientation of input fluctuations can result in autocorrelation increasing at the same time as variance is decreased. Insight into these phenomena can be gained by considering

the behavior of a simple low dimensional linear stochastic system. In the one-dimensional linear case of an Ornstein-Uhlenbeck process, $\mathrm{d}x = -ax\,\mathrm{d}t + b\,\mathrm{d}W$, the (normalized) autocorrelation is given by $\exp(-a\tau)$ and variance by $\frac{1}{2}b^2/a$, so both increase as the size of the eigenvalue a approaches zero. In particular, variance increases linearly with the variance of noise input b^2 (Gardiner, 2010). The same is true for a linearized two-dimensional system near a Hopf bifurcation (Steyn-Ross et al., 2006). If we naively assume that aligning input noise with the center eigenspace increases the amount of noise affecting the slow dynamical system of the center manifold, we would expect variance to be greater when the angle with the center eigenspace plane was smaller, which was not the case. From consideration of the normal form transformation (Roberts, 2008) it is rather the plane perpendicular to the stable eigenspace that should be relevant in determining the magnitude of noise driving the slow dynamics. Because the eigenvectors of the linearized Jansen-Rit system are far from orthogonal, that reference plane is almost normal to the center eigenspace plane, resulting in the observed reversal of the expected relationship between noise orientation and variance.

It may be possible to study these bifurcation indicators more specifically in a normal form model by considering a full center manifold reduction. Close to the bifurcation nonlinear terms can result in multiplicative noise in the slow dynamical system of the center manifold (Roberts, 2008). These occur in addition to the simpler additive noise that results directly from linear transformation of the input noise terms but so far we have yet to calculate the magnitude or importance of these multiplicative noise terms in the present system. Furthermore, any local analysis of the behavior close to the equilibrium point is valid only for the case of small noise, so that the state of the system remains local to the equilibrium point. That is not necessarily the case for this system, as suggested by larger output standard deviation near the bifurcation seen in the cases of H1 and H3p, which is comparable to the amplitude of the subsequent limit cycles. This implies that the system is exploring a wider region of phase space compared to the cases with high autocorrelation (H2 and H3u). Thus the structure of flow in the phase space further from the equilibrium point may be directly responsible for the quickly decaying autocorrelation in those cases. In particular if the center manifold curves away from the center eigenspace, then at a sufficient distance from the equilibrium point the directions of noise input which are “well aligned” and “poorly aligned” with the manifold may be reversed.

2.4.2 Power law scaling of output fluctuations

Analyzing the distributions of fluctuation sizes and durations, we observe the presence of a long power law scaling regime that extends over 4 orders of magnitude with a rapid truncation at the far right hand tail at the bifurcation when input fluctuations are normal to the reference plane. This power law scaling is not observed when input fluctuations have a significant projection onto the reference plane. Further away from the bifurcation, the power law regime extends for less than one order of magnitude so that the lengthy power law tail provides a signature of proximity to the bifurcation in that scenario.

A range of simple dynamical mechanisms are known to permit production of scale-free fluctuation structure of this kind. A relaxation process with a fractional operator formally yields a power law (Pareto) probability distribution of fluctuation durations (Sokolov and Klafter, 2005). Multiplicative noise (which arises when reducing oscillation dynamics of the model to two dimensions) can also in specific cases result in power law probability distributions (Anteneodo and Riera, 2005). However, the cause of the power law scaling of the distributions of fluctuations in our system is not yet determined.

2.5 Future work

This study considered autocorrelation in the output of a single Jansen-Rit model region, representing a small area of cortex of the order of 2–3 mm^2. For the question of potential detectability in EEG it remains to examine the effect on autocorrelation of combining the output of a large number of cortical regions, whose oscillations may be synchronized to a greater or lesser degree and where the output measurement function relating EEG to the combination of sources plays an important role.

Within the Jansen-Rit model we also observed indicators close to other bifurcation types, including switching between attractors in a bistable region near a cusp bifurcation and “flickering” or intermittent switching away from a stable fixed point in a monostable region near a sniper bifurcation, which are not explored further in this paper. Therefore it remains to examine the sensitivity of these and other indicators, such as mean switching times as bifurcations are approached, to noise orientation.

It is hoped that normal form analyses near the bifurcation will shed some light on the mechanism by which the input noise affects autocorrelation. A first step will be to examine a simpler normal form

system displaying the same behavior, where exact control over the shape of the center manifold can be afforded, initially targeting the limiting case of small fluctuations. Such an analysis will serve to separate the generic local effects of the Hopf bifurcation from global behavior due to excursion of the state further from the equilibrium point.

Examination of a normal form system will also be key to determining the reason for the power law scaling of fluctuation statistics. The results presented in this paper show that some of the indicators of instability reported in human EEG also arise in the output of a simple neural mass model near linear instability.

While similar indicators can also emerge from a critical phase transition in a complex, multi-scale system, we have shown in the present study that some of the same indicators can arise in a very different way, from the low-dimensional stochastic dynamics at a single scale: the mesoscopic scale of interacting populations. As the field advances, it will become increasingly important to move away from a single umbrella notion of "criticality" in brain dynamics towards defining a number of exact, and possibly distinct, mechanisms responsible for correlations and scale-free fluctuations in the time and/or spatial domains. It is certainly possible at this stage that multiple mechanisms play a role.

2.6 Supplementary figures

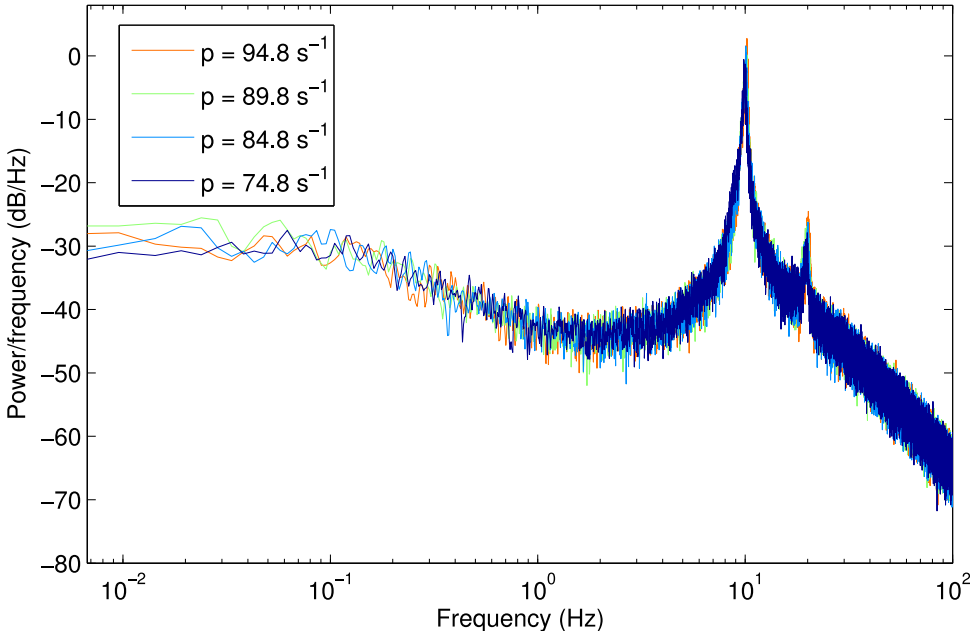


Figure 2.11: Power spectrum at H1, using a larger window size of 150 s (750000 samples) to show lower frequencies from $6.7 \times 10^{-3}\,Hz$.

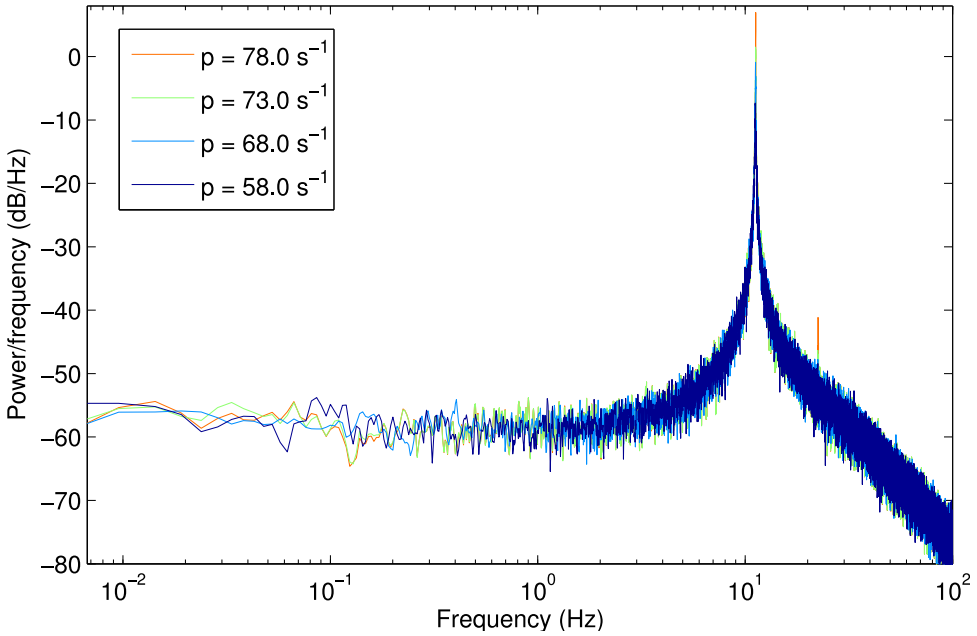


Figure 2.12: Power spectrum at H2, using a larger window size of 150 s (750000 samples) to show lower frequencies from $6.7 \times 10^{-3}\,Hz$.

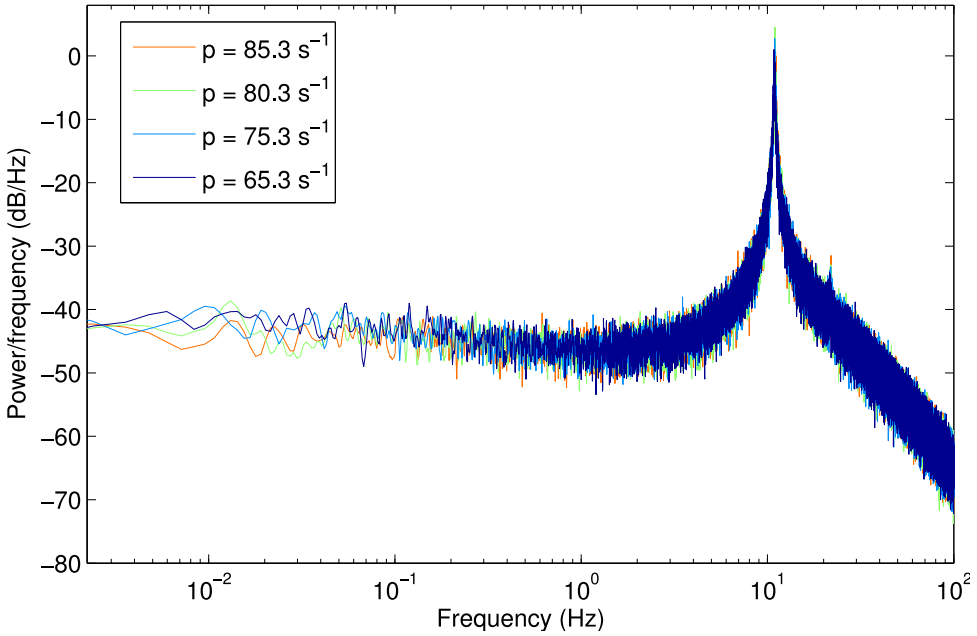


Figure 2.13: Power spectrum for scenario H3p, using a larger window size of 450 s (2250000 samples) to show lower frequencies from $2.2 \times 10^{-3}\ Hz$.

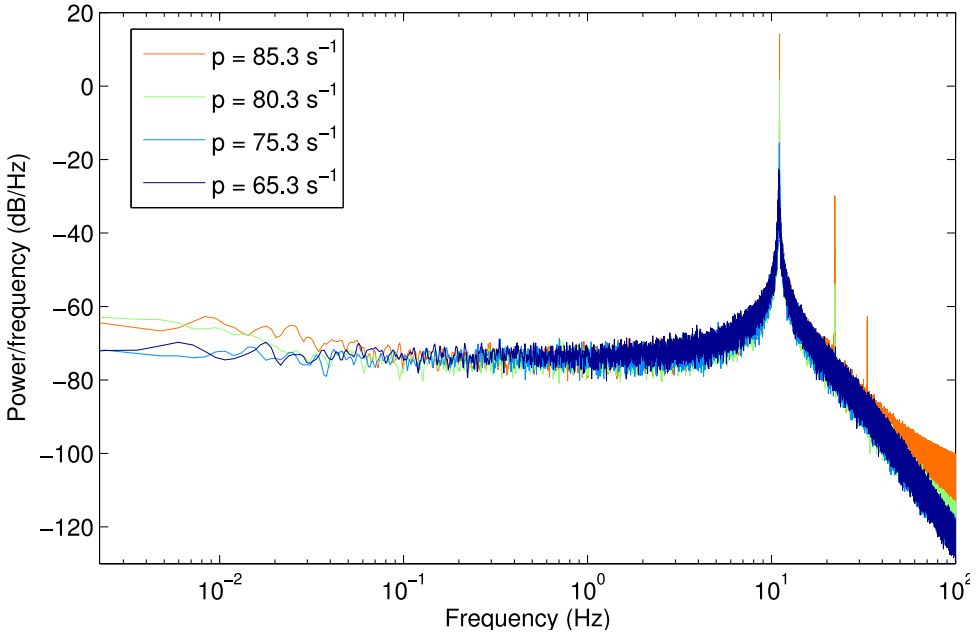


Figure 2.14: Power spectrum for scenario H3u, using a larger window size of 450 s (2250000 samples) to show lower frequencies from $2.2 \times 10^{-3}\ Hz$.

Chapter 3

Using normal form transformations to parameterize minimal biological models near bifurcation

In dynamical models of biological or physical systems the underlying mechanism behind a transition in behavior is often a local bifurcation of equilibria. In these cases a normal form model is then the simplest mathematical model that distills the mechanism, allowing one to test the hypothesis that the given bifurcation is sufficient to account for the observed behavior in the real system. In practice, somewhat arbitrary parameter values are sometimes used in such normal form models, the assumption being that the evolution in phase space will be qualitatively the same. But this assumption breaks down when the systems are not isolated. Instead, we show how a normal form transformation can constrain the parameters of a normal form model to match the original biophysical model, giving an explicit mapping from the full model near bifurcation to a local, minimal model, and making the coupling function between normal form models consistent with the biophysical coupling model of the full system. We provide a Mathematica package to derive normal form transformations for arbitrary systems near local bifurcations. As an application, we demonstrate that this approach is useful when modeling the synchronization of nonlinear oscillators near a Hopf bifurcation. If the parameters of the topological normal form are used with symmetrical diffusive coupling the synchronization behavior of a minimal model does not match the original system. But when a Poincaré normal form transformation

is used to parameterize the minimal model and coupling, the synchronization behavior matches the original system.

3.1 Background

3.1.1 Normal form models in neuroscience

There is a long history of simplified abstract models being used in computational neuroscience to understand synchronization of neural activity and responses to input, at all scales from single neuron to mesoscopic and macroscopic brain network scale. The utility of these minimal oscillator models was discussed in Section 1.4. As introduced in Section 1.7.2, each type of bifurcation has a corresponding topological normal form system, which is the "simplest" system of equations exhibiting that type of bifurcation.

In neuroscience this normal form system itself may be used as a model when it is hypothesized that the phenomenon of interest is explained by a particular bifurcation occurring in the real biological system. In this situation the maximal simplicity of the normal form model in capturing the mechanism gives it greater explanatory power than a detailed biological model.

A second reason for using a normal form model is that the simple form may be analytically tractable. Existing theorems and mathematical results may be directly applicable to the reduced system. This is particularly true in the study of oscillator synchronization, where many existing mathematical results are available.

A third reason for using a normal form system as a model is that it requires less computational resources to simulate, thus making larger scale simulations of coupled networks more feasible (Goodfellow and Glendinning, 2013). This is a reason of lesser importance.

In the case of the Hopf bifurcation, its two-dimensional normal form model is so ubiquitous that it is sometimes given a separate name: the Landau-Stuart equation (Pikovsky et al., 2003).

Modeling with normal forms

For neuroscience applications a normal form model may be arrived at in several ways:

A detailed biological model may be developed first, then a bifurcation analysis of this model indicates a specific type of bifurcation which is responsible for a transition occurring at the boundary of a region in parameter space. This transition can then be modeled separately by a normal form system of equations for that bifurcation type.

Alternatively specific features of the dynamics may be desired to be reproduced with a minimal model. For example Goodfellow et al. (2011) used an 8-dimensional Jansen-Rit neural mass model which exhibited intermittent transition between low- and high-amplitude oscillations. Having identified these dynamical features the authors later chose a 2-dimensional 7th order generalized Hopf normal form model in Goodfellow and Glendinning (2013) to implement this transition, as it allows bistability between an inner and an outer limit cycle. (This example shows how normal form models may be used where bistability rather than a bifurcation is the mechanism of transition.)

Finally, phenomenological models may be constructed from a repertoire of simple normal forms based on observed time series properties. An excellent example of this approach is given by Jirsa et al. (2014).

In some cases it is possible first to derive a reduction of an oscillating biophysical model to the Poincaré normal form beyond a Hopf bifurcation, then to continue further and perform a reduction to a one-dimensional phase-oscillator model from the normal form model (Schuster and Wagner, 1990; Brown et al., 2004; Daffertshofer and van Wijk, 2011; Ashwin and Rodrigues, 2016). This is valuable for mathematical analysis of the synchronization dynamics, as discussed in Section 1.4.1, but is only possible if the oscillator parameters remain away from the bifurcation point.

Bifurcations, local and global, that can be the mechanism of oscillation onset and amenable to a minimal model include supercritical Hopf, SNIPER (saddle-node in a periodic orbit), saddle-node of limit cycles, homoclinic bifurcation (Guckenheimer and Holmes, 1985). At the single-neuron scale these bifurcation types have all appeared in various mathematical models in neuroscience (Brown et al., 2004; Izhikevich, 2007).

An example of the application of normal form models to large scale network synchronization in epilepsy research is Petkov et al. (2014). This modeled local bistable oscillatory dynamics with Generalized Hopf normal form models, (as discussed in the context of Figure 1.3), with network structures defined by measured time series of patients and healthy controls.

3.1.2 Importance of phase-amplitude oscillator parameters

An ϵ-dependent Poincaré normal form model of oscillations near a supercritical Hopf bifurcation has the form

$$\begin{aligned} \dot{r} &= \gamma_\epsilon \epsilon r - a r^3 \\ \dot{\theta} &= \omega + \omega_\epsilon \epsilon + b r^2. \end{aligned} \tag{3.1}$$

This can be compared to the maximally simple topological normal form, a special case of Eq. (3.1) with parameters fixed:

$$\begin{aligned} \dot{r} &= \epsilon r - r^3 \\ \dot{\theta} &= 1. \end{aligned} \tag{3.2}$$

With arbitrary parameters Eqs (3.1) and (3.2) describe transitions to oscillation that are qualitatively the same: their flows in state space are topologically equivalent, as discussed in Section 1.7.2. However when oscillators are coupled together as a network, previous work has shown that besides the natural frequencies ω the parameters b and γ_ϵ in the Poincaré normal form will also have a significant effect on the collective dynamics of synchronization.

Aronson et al. (1990) examined in detail the case of two coupled oscillators near Hopf bifurcation and showed through bifurcation analysis of the 4-dimensional coupled system that the shear parameter b affects not only the thresholds where synchronization appears, but also the qualitative types of synchronous behavior that are possible, as shown in Figure 3.1 reproduced from Aronson et al. (1990). Regions of bistability between phase locking and phase drift or bistability between phase locking and oscillator death also occur in the case of nonzero b.

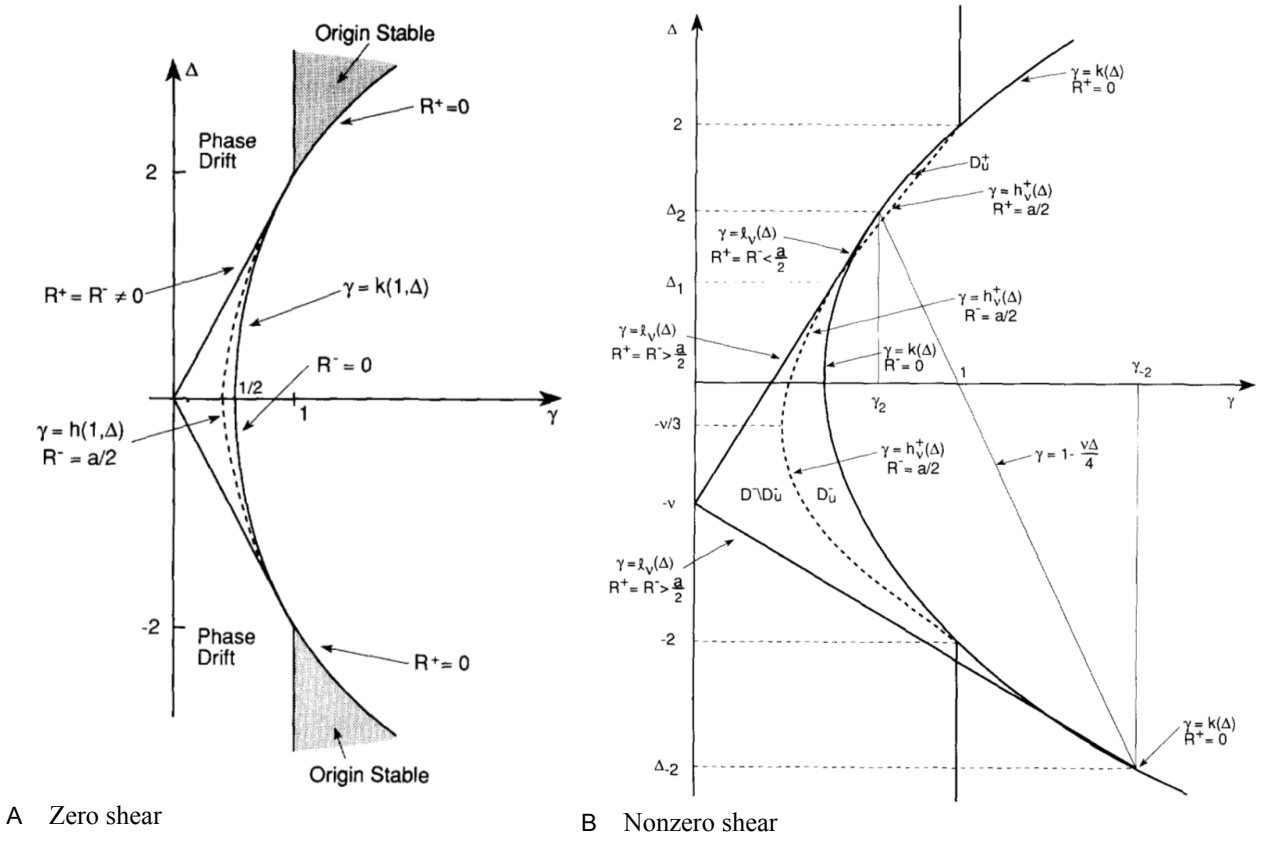


Figure 3.1: A nonzero shear parameter b in the Hopf normal form breaks the symmetry of the bifurcation diagram for the coupled system of two oscillators. Shown are the bifurcation curves that apply at the point of amplitude symmetry ($r_1 = r_2$), as a function of parameter γ (proportional to coupling strength) and the difference $\Delta = (\omega_2 - \omega_1) + (b_2 - b_1)$ between oscillators. Reprinted from Aronson et al. (1990) with permission from Elsevier.

Daffertshofer and van Wijk (2011) demonstrated that accurate amplitude dynamics are essential to modeling synchronization behavior in networks. They treated a network of coupled Wilson and Cowan (1972) neural mass models, using averaging to reduce the limit cycle of each to a phase oscillator model, whilst retaining the dependence of the coupling strength between phase oscillators on the source and target oscillation amplitudes in the original model. In this way they derived the dependence of phase coupling on oscillation amplitude. This was supported by numerical simulations with different amplitude distributions and network topologies.

Therefore, close to a supercritical Hopf bifurcation where the amplitude is (proportionally) highly variable, it is expected that the natural frequencies ω, varying amplitudes (governed by parameters γ_ϵ and a) and shear b all affect network synchronization. In this regime the oscillation amplitude is not linearly stable, so the requirements for an approximate phase reduction to one dimension are in general not satisfied: the dynamics are 2-dimensional in character (Aronson et al., 1990). The exception is

the regime further beyond bifurcation where the bifurcation parameter ϵ is kept much larger than the coupling strength c (and much larger than noise intensity σ if the amplitude is also perturbed by noise), in which case a reduction to coupled phase oscillators is possible (Ashwin and Rodrigues, 2016).

3.1.3 Normal form transformations

Normal form transformations were introduced in Section 1.7.2. For local bifurcations the theory of normal forms includes methods to find explicitly a smooth mapping from a complicated system near bifurcation to its normal form. In the following Sections 3.2 and 3.3 we review one such method. The iterative method applied in this paper for equivariant normal forms is originally due to Poincaré (1879) and a more general theory subsequently developed (Murdock, 2003; Guckenheimer and Holmes, 1985; Kuznetsov, 2010; Golubitsky et al., 1988). An outline of the historical development is given in Murdock (2003).

3.2 Example system

To introduce the method of Poincaré that will be applied to biophysical models, we first illustrate the steps with a minimal 3-dimensional example:

$$\begin{aligned}\dot{x}_1 &= \epsilon x_1 - 9x_2 - x_1x_3\\ \dot{x}_2 &= 9x_1 + \epsilon x_2 - x_2x_3\\ \dot{x}_3 &= -2x_3 + 4(x_1^2 + x_2^2).\end{aligned} \tag{3.3}$$

In practical application, such as the neural mass model example in Section 3.5 below, the starting biophysical model will have higher dimension and so the reduction is best done automatically in a computer algebra system, such as the Mathematica package accompanying this paper, or the Reduce implementation of Roberts (2009). This first example is small enough to write all steps explicitly, making the method clear.

Briefly examining the dynamics of Eq. (3.3), from the equations it can be seen that the origin is a fixed point, and that the x_3 direction is stable having eigenvalue -2 (from the coefficients of the

linear terms), so that the state collapses exponentially to a paraboloid-shaped center manifold surface $x_3 = 2(x_1^2 + x_2^2)$ (determined by $\dot{x}_3 = 0$). It is on this curved surface that a stable limit cycle emerges as ϵ is increased, given by $x_3 = \epsilon$, $x_1^2 + x_2^2 = \epsilon/2$.

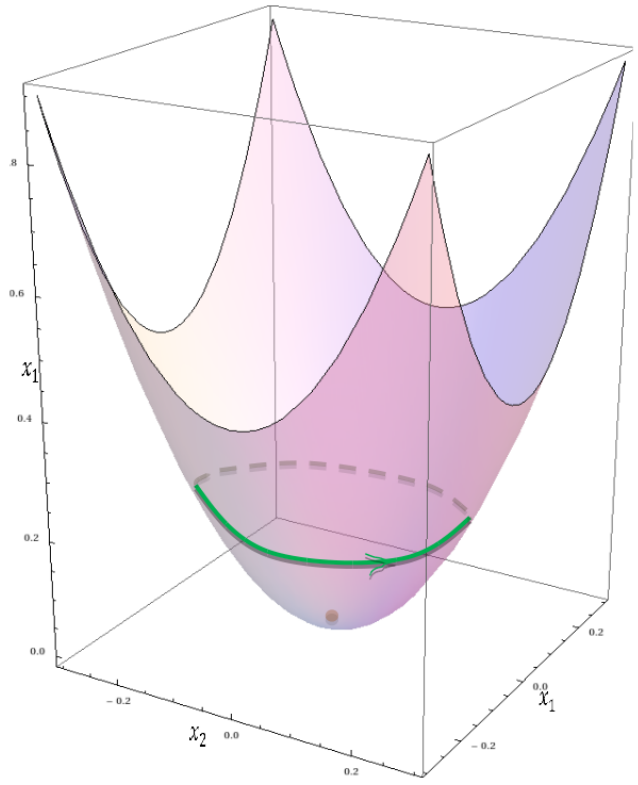


Figure 3.2: The example system of Eq. (3.3) showing the paraboloid shape of its center manifold, with a limit cycle of stable oscillations (green) emerging on this surface beyond the bifurcation point.

The first step is to translate the bifurcating equilibrium point to the origin and apply an initial linear transformation to the system so that eigenspaces at the bifurcation point are orthogonal and aligned with coordinate axes. In this example the bifurcation point is already at the origin and eigenspaces already aligned with the coordinate axes, so no action is needed: for this example the initial linear transformation is the identity.

Next the system right-hand-side is approximated by a truncated Taylor series about the origin. For this example again no action is needed as the right-hand-side is already a polynomial vector field.

Now the normal form algorithm can be applied to find a smooth change of variables which maps the deterministic system to its normal form. Eq. (3.3) can be written as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = A(\epsilon) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} -x_1 x_3 \\ -x_2 x_3 \\ 4x_1^2 + 4x_2^2 \end{bmatrix}, \tag{3.4}$$

where

$$A(\epsilon) = \begin{bmatrix} \epsilon & -9 & 0 \\ 9 & \epsilon & 0 \\ 0 & 0 & -2 \end{bmatrix} \quad \text{so at the bifurcation point} \quad A = \begin{bmatrix} 0 & -9 & 0 \\ 9 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix}. \tag{3.5}$$

The right-hand-side of our system 3.4, which we want to simplify, defines a flow on the phase space $\mathbb{R}^3$. Because we approximated it with a Taylor series, this flow vector field is an element of the space $\mathcal{V}$ of all three-dimensional polynomial vector fields in three variables.

The first iteration of the algorithm will simplify second order terms, operating in the smaller subspace $\mathcal{V}_2$ that consists only of polynomial vector fields of homogeneous order 2, that is the space of all possible second order terms. Explicitly, for a system of three variables $\mathcal{V}_2$ is 18-dimensional, with standard ordered basis

$$\left(\begin{bmatrix} x_1^2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} x_1 x_2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} x_1 x_3 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} x_2^2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} x_2 x_3 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} x_3^2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ x_1^2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ x_1 x_2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ x_1 x_3 \\ 0 \end{bmatrix}, \right.$$
$$\left. \begin{bmatrix} 0 \\ x_2^2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ x_2 x_3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ x_3^2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ x_1^2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ x_1 x_2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ x_1 x_3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ x_2^2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ x_2 x_3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ x_3^2 \end{bmatrix} \right). \tag{3.6}$$

Using this ordered basis the system's existing second order terms,

$$\mathbf{f}_{\{2\}}(\mathbf{x}) = \begin{bmatrix} -x_1 x_3 \\ -x_2 x_3 \\ 4x_1^2 + 4x_2^2 \end{bmatrix}, \tag{3.7}$$

correspond to the following column vector in $\mathbb{R}^{18}$:

$$\hat{\mathbf{f}}_{\{2\}} = (0, 0, -1, 0, 0, 0, 0, 0, 0, 0, -1, 0, 4, 0, 0, 4, 0, 0)^T. \tag{3.8}$$

As will be discussed in general in Section 3.3.2, the normal form algorithm finds a simplifying transformation by using the linear operator $\mathfrak{L}_A \colon \mathcal{V} \to \mathcal{V}$ defined by

$$\mathfrak{L}_A \mathbf{q}(\mathbf{x}) = (D\mathbf{q}(\mathbf{x}))A\mathbf{x} - A\mathbf{q}(\mathbf{x}) \quad \text{for any } \mathbf{q}(\mathbf{x}) \in \mathcal{V}. \tag{3.9}$$

To simplify the system's second order terms what is important is the action of $\mathfrak{L}_A$ on the subspace of second order terms $\mathcal{V}_2$. The matrix representation of the action of $\mathfrak{L}_A$ on $\mathcal{V}_2$ can be seen explicitly by applying the definition Eq. (3.9) to the basis elements of $\mathcal{V}_2$, giving

$$\hat{L}_{A\{2\}} = \begin{bmatrix}
0 & 9 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-18 & 0 & 0 & 18 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -2 & 0 & 9 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -9 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -4 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 \\
-9 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -9 & 0 & 0 & 0 & 0 & -18 & 0 & 0 & 18 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -9 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -9 & 0 & 0 & 0 & -9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -9 & 0 & 0 & 0 & -9 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -9 & 0 & 0 & 0 & 0 & 0 & -4 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 9 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -18 & 2 & 0 & 18 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -9 & 0 & 2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -9 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2
\end{bmatrix}.$$

In this example system this matrix has full rank, i.e. the image of $\mathcal{V}_2$ under $\mathfrak{L}_A$ is the whole subspace $\mathcal{V}_2$. This indicates that all second order terms can be removed by a suitable change of variables,

$$\mathbf{x} = \mathbf{u} + \mathbf{q}_{\{2\}}(\mathbf{u}), \tag{3.10}$$

so that the second order terms after this transformation will be $\mathbf{g}_{\{2\}}(\mathbf{u}) = \mathbf{0}$, corresponding to the zero vector in $\mathbb{R}^{18}$:

$$\hat{\mathbf{g}}_{\{2\}} = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)^T. \tag{3.11}$$

The key result of normal form theory is that the required transformation $\mathbf{q}_{\{2\}}(\mathbf{u})$ will always satisfy the equation

$$\mathfrak{L}_A \mathbf{q}_{\{2\}}(\mathbf{x}) = \mathbf{f}_{\{2\}}(\mathbf{x}) - \mathbf{g}_{\{2\}}(\mathbf{x}). \tag{3.12}$$

So to find $\mathbf{q}_{\{2\}}(\mathbf{u})$ explicitly, we need only solve the corresponding linear system in $\mathbb{R}^{18}$,

$$\hat{L}_{A\{2\}}\hat{\mathbf{q}}_{\{2\}} = \hat{\mathbf{f}}_{\{2\}} - \hat{\mathbf{g}}_{\{2\}}. \tag{3.13}$$

This gives

$$\begin{aligned}\hat{\mathbf{q}}_{\{2\}} &= \hat{L}^{-1}_{A\{2\}}\hat{\mathbf{f}}_{\{2\}} \\ &= (0, 0, \tfrac{1}{2}, 0, 0, 0, 0, 0, 0, 0, \tfrac{1}{2}, 0, 2, 0, 0, 2, 0, 0)^T,\end{aligned} \tag{3.14}$$

which in the original space $\mathcal{V}$ of vector fields corresponds to

$$\mathbf{q}_{\{2\}}(\mathbf{u}) = \begin{bmatrix} \frac{1}{2}u_1u_3 \\ \frac{1}{2}u_2u_3 \\ 2u_1^2 + 2u_2^2 \end{bmatrix}. \tag{3.15}$$

Indeed, applying the newly found transformation (3.10) to Eq. (3.4) removes all quadratic terms as desired, introducing cubic terms as a side effect:

$$\begin{bmatrix} \dot{u}_1 \\ \dot{u}_2 \\ \dot{u}_3 \end{bmatrix} = A \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + \begin{bmatrix} -2u_1^3 - 2u_1u_2^2 - u_1u_3^2/2 \\ -2u_1^2u_2 - 2u_2^3 - u_2u_3^2/2 \\ 4u_1^2u_3 + 4u_2^2u_3 \end{bmatrix} + O(|\mathbf{u}|^4). \tag{3.16}$$

The second iteration of the algorithm will now simplify cubic terms, while leaving first and second order terms unchanged. Working in the subspace $\mathcal{V}_3$, we seek a further change of variables, this time equal to the identity up to second order,

$$\mathbf{u} = \mathbf{y} + \mathbf{q}_{\{3\}}(\mathbf{y}). \tag{3.17}$$

For a system of three variables $\mathcal{V}_3$ is 30-dimensional. A basis for $\mathcal{V}_3$ is chosen analogously to Eq. (3.6). Considering the cubic terms of Eq. (3.16) as a vector $\mathbf{f}_{\{3\}}(\mathbf{u})$ in the space $\mathcal{V}_3$, this cubic field corresponds to a column vector $\hat{\mathbf{f}}_{\{3\}}$ in $\mathbb{R}^{30}$,

$$\hat{\mathbf{f}}_{\{3\}} = (-2, 0, 0, -2, 0, -\tfrac{1}{2}, 0, 0, 0, 0, 0, -2, 0, 0, 0, 0, -2, 0, -\tfrac{1}{2}, 0, 0, 0, 4, 0, 0, 0, 0, 4, 0, 0)^T. \tag{3.18}$$

Similarly to the first iteration, the action of $\mathfrak{L}_A$ on $\mathcal{V}_3$ is represented by a matrix $\hat{L}_{A\{3\}} \in \mathbb{M}_{30}(\mathbb{R})$,

$$\hat{L}_{A\{3\}} = \left[\begin{array}{cccccccccccccccccccccccccccccc}
0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-27 & 0 & 0 & 18 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -2 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -18 & 0 & 0 & 0 & 0 & 27 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -18 & 0 & -2 & 0 & 0 & 18 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -4 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -9 & 0 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -9 & 0 & 0 & -4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -27 & 0 & 0 & 18 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -18 & 0 & 0 & 0 & 0 & 27 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -18 & 0 & -2 & 0 & 0 & 18 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -4 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -9 & 0 & 0 & 0 & 0 & 0 & 0 & -9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -9 & 0 & 0 & 0 & 0 & 0 & 0 & -9 & 0 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -9 & 0 & 0 & 0 & 0 & 0 & 0 & -9 & 0 & 0 & -4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -27 & 2 & 0 & 18 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -18 & 0 & 2 & 0 & 0 & 27 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -18 & 0 & 0 & 0 & 0 & 18 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & 0 & 9 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -9 & 0 & 0 & 2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -9 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -9 & 0 & 0 & -2 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -4
\end{array}\right]$$

This time the matrix is not invertible, so bases for the image and kernel of $\hat{L}_{A\{3\}}$ are computed, having dimension 27 and 3 respectively, and the vector $\hat{\mathbf{g}}_{\{3\}}$ is found as the oblique projection of the vector $\hat{\mathbf{f}}_{\{3\}}$ along the subspace $\operatorname{im} \hat{L}_{A\{3\}}$ into the subspace $\ker \hat{L}_{A\{3\}}$,

$$\hat{\mathbf{g}}_{\{3\}} = (-2,0,0,-2,0,0,0,0,0,0,0,-2,0,0,0,0,-2,0,0,0,0,0,0,4,0,0,0,0,0,4,0,0)^T, \tag{3.19}$$

which corresponds to the cubic vector field

$$\mathbf{g}_{\{3\}}(\mathbf{y}) = \begin{bmatrix} -2y_1^3 - 2y_1y_2^2 \\ -2y_1^2y2 - 2y_2^3 \\ 4y_1^2y_3 + 4y_2^2y_3 \end{bmatrix}. \tag{3.20}$$

These will be the simplified cubic terms of the normal form. To find the required transformation $\mathbf{q}_{\{3\}}(\mathbf{y})$ a linear equation, Eq. (3.30), is again solved, by solving the corresponding linear system in $\mathbb{R}^{30}$,

$$\hat{L}_{A\{3\}}\hat{\mathbf{q}}_{\{3\}} = \hat{\mathbf{f}}_{\{3\}} - \hat{\mathbf{g}}_{\{3\}}, \tag{3.21}$$

which gives

$$\hat{\mathbf{q}}_{\{3\}} = (0,0,0,0,0,\tfrac{1}{8},0,0,0,0,0,0,0,0,0,0,0,0,\tfrac{1}{8},0,0,0,0,0,0,0,0,0,0,0)^T, \tag{3.22}$$

corresponding to

$$\mathbf{q}_{\{3\}}(\mathbf{y}) = \begin{bmatrix} y_1 y_3^2/8 \\ y_2 y_3^2/8 \\ 0 \end{bmatrix}. \tag{3.23}$$

Applying the transformation 3.17 to Eq. (3.16) indeed gives the cubic terms of 3.20. Thus we have computed the normal form to third order:

$$\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \end{bmatrix} = \begin{bmatrix} \epsilon y_1 - 9y_2 - 2y_1^3 - 2y_1 y_2^2 \\ 9y_1 + \epsilon y_2 - 2y_1^2 y_2 - 2y_2^3 \\ -2y_3 + 4y_1^2 y_3 + 4y_2^2 y_3 \end{bmatrix} + O(|\mathbf{y}|^4). \tag{3.24}$$

In this normal form system (3.24) the equations for y_1 and y_2 now no longer depend on the stable variable y_3. Writing the (y_1, y_2) equations in polar form,

$$\begin{aligned} \dot{r} &= \epsilon r - 2r^3 \\ \dot{\theta} &= 9. \end{aligned} \tag{3.25}$$

This is a Poincaré normal form for a supercritical Hopf bifurcation. Thus we have a simpler 2-dimensional system that captures the emerging oscillation behavior of the original 3-dimensional system. If desired the transformation that maps the original system to the minimal normal form system can be found explicitly by composing the transformations from the two iterations.

3.3 Review of normal form transformation method

Having studied a minimal example, we briefly review the basic theory behind this method, to understand how these steps work in the general case. For a clear and comprehensive exposition of normal form theory, including its generalizations and more sophisticated algorithms, see Murdock (2003). The simplest method, presented here, has been automated in the accompanying Mathematica package (Appendix D) which can be given an arbitrary system near a local bifurcation and will output both the transformation and the resulting simplified system.

3.3.1 Initial translation and linear transformation

After a local bifurcation point has been located in the original system (for example by using the numerical continuation software MATCONT, Dhooge et al. (2006)) a translation is applied to shift the bifurcation to the origin of phase space and parameter space.

The normal form transformation will produce a simplified system by a change of variables that imposes on the non-linear terms the same symmetry possessed by the linear term (i.e. the Jacobian matrix). Therefore a preparatory step is to make that linear term as simple as possible, by linearly transforming the system so that its Jacobian is in Jordan real form. This aligns the coordinate system with (generalized) eigenspaces of the Jacobian. As we are at a local bifurcation point, these eigenspaces include a center eigenspace of dimension $c \geq 1$ (depending on bifurcation type) tangent to the center manifolds of the bifurcation (Guckenheimer and Holmes, 1985; Kuznetsov, 2010). Thus after linear transformation the center eigenspace of the bifurcation will correspond to c specific variables, but in general at this stage these variables will still be coupled nonlinearly to the remaining (stable and unstable) variables, because the center manifolds are curved. A geometric effect of the subsequent transformations will be to flatten the center manifolds so they coincide with the flat center eigenspace up to mth order.

Assuming derivatives exist at the origin the system is now approximated locally by Taylor expansion. To compute the normal form to mth order it is sufficient to retain up to mth order terms in the Taylor approximation. In this way the flow $\mathbf{f}(\mathbf{x})$ near the origin is expressed as a polynomial vector field with each of the d components being a polynomial in variables $x_1, \cdots, x_d$.

3.3.2 Near identity transformation

To complete the normal form transformation the aim is to find a smooth invertible transformation,

$$\mathbf{x} = \mathbf{y} + \mathbf{q}(\mathbf{y}), \tag{3.26}$$

which maps the deterministic system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, 0)$ at the bifurcation point to a normal form system $\dot{\mathbf{y}} = \mathbf{g}(\mathbf{y})$ where the flow of the center variables are decoupled and have the symmetry (equivariance)

of the linear term. We seek a transformation that is *near-identity* in the sense that it equals the identity transformation to first order, affecting only nonlinear terms in the Taylor approximation, that is $|\mathbf{q}(\mathbf{y})| = O(|\mathbf{y}|^2)$.

A polynomial is called *homogenous of degree* m if each of its terms has indices that sum to m, for example $x_1^2 x_2 + x_2^3$ is homogenous of degree 3. The set of all d-dimensional polynomial vector fields in d variables forms a vector space $\mathcal{V}$, with a standard basis given by monomial vector fields of the form

$$\mathbf{e}_j \prod_{i=1}^{d} y_i^{n_i}, \tag{3.27}$$

where $\mathbf{e}_j$ are the standard unit basis vectors of $\mathbb{R}^d$. The space $\mathcal{V}$ will be viewed as the direct sum of subspaces $\mathcal{V}_m$ where an element of $\mathcal{V}_m$ has as each of its components either homogenous polynomials of degree m, or zero. Each transformation that we use will be an endomorphism on these subspaces $\mathcal{V}_m$, so will always map a vector field of homogenous degree m either to a field homogenous of the same degree m in the new variables, or to zero.

A key result is that the terms of order m in the deterministic system which can be removed by a transformation such as Eq. (3.26) are completely determined by the linear part of the system, $A = D\mathbf{f}(\mathbf{x})|_{\mathbf{x}=\mathbf{0}}$ via the action of the linear operator $\mathfrak{L}_A \colon \mathcal{V} \to \mathcal{V}$ defined by

$$\mathfrak{L}_A \mathbf{q}(\mathbf{x}) = [\mathbf{q}(\mathbf{x}), A\mathbf{x}] = (D\mathbf{q}(\mathbf{x}))A\mathbf{x} - A\mathbf{q}(\mathbf{x}). \tag{3.28}$$

(Recall from Section 1.7.2 that the whole system is in normal form if $\mathfrak{L}_A \mathbf{f}(\mathbf{x}) = 0$).
The image of each subspace $\mathcal{V}_m$ under this operator $\mathfrak{L}_A$ gives the entire set of order m terms that may arise in the transformed system from any possible polynomial transformation $\mathbf{q}$. So if there are terms in the original system that are not in the image of $\mathfrak{L}_A$ then these cannot be eliminated by any such transformation. Thus it is terms in a subspace complementary to the image of $\mathfrak{L}_A$ that will form the new, simplified system. In this way the symmetry that can be achieved in the normal form is determined by the linear part A alone. In particular, expressing A as the sum of a semisimple and a nilpotent matrix $A = S + N$, all terms in the transformed system will be in the kernel of $\mathfrak{L}_S$. This implies all terms commute with $S\mathbf{x}$ as vector fields, so that the new system is equivariant under the flow $e^{\alpha S}$ generated by matrix S (Murdock, 2003). For our application to Hopf bifurcations this ensures

that these systems will be mapped by a polynomial transformation to a normal form that has circular symmetry.

There remains the freedom to select a subspace complementary to the image of $\mathfrak{L}_A$. In cases such as Section 3.2, where the representation of $\mathfrak{L}_A$ on $\mathcal{V}_m$ is a semisimple matrix, its kernel can be chosen as the complement, giving the semisimple normal form style. Otherwise, as in the non-semisimple example of the Jansen-Rit model in Section 3.5 below, the kernel of the representation of $\mathfrak{L}_{A^\dagger}$ on $\mathcal{V}_m$ is used instead, giving the inner product normal form style. (Murdock, 2003)

To compute the required transformation the simplest algorithm, which we use here, is iterative. The first iteration derives a transformation $\mathbf{x} = \mathbf{u} + \mathbf{q}_{\{2\}}(\mathbf{u})$ to simplify only the second-order terms (also modifying third and higher order terms as a side effect). After choosing a basis for the space $\mathcal{V}_2$ the existing second-order terms $\mathbf{f}_{\{2\}}(\mathbf{x})$ are considered as a vector in $\mathcal{V}_2$. The desired simplified second-order terms $\mathbf{g}_{\{2\}}(\mathbf{x})$ are found by explicitly calculating the matrix representing the action of $\mathfrak{L}_A$ on the subspace $\mathcal{V}_2$ then computing the oblique projection of the vector $\mathbf{f}_{\{2\}}(\mathbf{x})$ along the image of $\mathfrak{L}_A$ into the chosen subspace complementary to the image of $\mathfrak{L}_A$. This choice of $\mathbf{g}_{\{2\}}$ ensures that $\mathbf{f}_{\{2\}}(\mathbf{x}) - \mathbf{g}_{\{2\}}(\mathbf{x})$ lies within the image of $\mathfrak{L}_A$, so that there exists $\mathbf{q}_{\{2\}}(\mathbf{x}) \in \mathcal{V}_2$ such that

$$\mathfrak{L}_A \mathbf{q}_{\{2\}}(\mathbf{x}) = \mathbf{f}_{\{2\}}(\mathbf{x}) - \mathbf{g}_{\{2\}}(\mathbf{x}). \tag{3.29}$$

Thus the required transformation $\mathbf{q}_{\{2\}}$ to simplify second order terms can be found as the pre-image of $\mathbf{f}_{\{2\}}(\mathbf{x}) - \mathbf{g}_{\{2\}}(\mathbf{x})$ by solving Eq. (3.29) as a linear system in $\mathcal{V}_2$.

The second iteration then proceeds similarly, leaving first and second order terms unchanged and simplifying third order terms, by considering those terms as a vector in the subspace $\mathcal{V}_3$ and projecting that vector into a subspace complementary to the image of $\mathfrak{L}_A$ acting on $\mathcal{V}_3$. At each iteration a similar *homological equation*,

$$\mathfrak{L}_A \mathbf{q}_{\{m\}}(\mathbf{x}) = \mathbf{f}_{\{m\}}(\mathbf{x}) - \mathbf{g}_{\{m\}}(\mathbf{x}), \tag{3.30}$$

is solved to find the incremental transformation $\mathbf{q}_{\{m\}}(\mathbf{x})$ which simplifies terms of order m. After $n-1$ iterations, the system is in normal form up to order n.

In both the semisimple and non-semisimple cases, the constraints satisfied by the kernel of $\mathfrak{L}_A$ ensure that the transformation flattens any curved center manifold of the original system, so that in the new

coordinates the center manifold coincides with the center eigenspace (to nth order) (Murdock, 2003). This implies that in the transformed system on the center manifold the center variables undergoing bifurcation are decoupled from the others, so that the bifurcation dynamics of interest can be analyzed as a separate c-dimensional subsystem.

This algorithm also gives explicitly the smooth coordinate transformation that maps the original system to its normal form, by composing the incremental transformations used in each iteration. Thus, where the normal form is computed to nth order the overall transformation is also expressed as a nth-order polynomial vector field.

3.3.3 Parameter-dependent transformations

In Section 3.3.2 we first set $\epsilon = 0$ to find a transformation based on the system exactly at the bifurcation point. The result was a single transformation which at bifurcation locally maps the system to its normal form. Having found this transformation, it can then be used to transform the original perturbed system including bifurcation parameter ϵ. But in general the transformed system for $\epsilon \neq 0$ will then contain ϵ-dependent terms that lack the symmetries of the normal form (i.e. equivariance under the flow generated by the Jacobian). As ϵ is varied to take the system through the bifurcation the transformed system is only approximately in normal form for $\epsilon \neq 0$.

As an improvement, we can promote the bifurcation parameter ϵ to the status of a dynamical variable, forming the $n + 1$ dimensional extended system,

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \epsilon) \\ \dot{\epsilon} &= 0, \end{aligned} \tag{3.31}$$

which now has a $c + 1$ dimensional center manifold. The normal form algorithm of Section 3.3.2 can then be applied to this extended system, yielding an ϵ-dependent transformation that leaves the $\dot{\epsilon} = 0$ equation unchanged (Guckenheimer and Holmes, 1985; Kuznetsov, 2010). For computer algebra implementation with truncated power series, we first rescale bifurcation parameters so that variables of the extended system are of the same order when truncating power series. That is, for each bifurcation parameter ϵ we extend the system by introducing a new variable $\alpha = \alpha(\epsilon)$ with $\dot{\alpha} = 0$ and $O(\alpha) =$

$O(x_i)$ near the bifurcation point. The result, on returning to n dimensions, is an ϵ-dependent change of variables,

$$\mathbf{x} = \mathbf{y} + \mathbf{q}(\mathbf{y}, \epsilon), \tag{3.32}$$

instead of Eq. (3.26), and an n-dimensional normal form system that now remains in normal form as ϵ is varied to take the system through the bifurcation. As before this contains a c-dimensional center manifold subsystem decoupled from the other variables, giving a reduced model. By using the extended system this reduced model is simpler, and in systems with further perturbations such as noise or coupling from another system it more accurately captures the effect of changing ϵ on the center manifold dynamics. (See for example Appendix B.)

3.4 Supercritical Hopf bifurcations

The preceding sections (3.3.1 - 3.3.3) and the accompanying Mathematica package apply to any local bifurcations of equilibria, including those of higher codimension. In this section we turn to specific considerations for the supercritical Hopf normal form that will appear in the application below.

3.4.1 Rescaling of oscillation amplitude

Applying the methods of Section 3.3.2 or Section 3.3.3, there is still one degree of freedom remaining in selecting a normal form. In the case of the Hopf bifurcation normal form,

$$\begin{aligned} \dot{r} &= \epsilon r - a r^3 \\ \dot{\theta} &= \omega + b r^2, \end{aligned} \tag{3.33}$$

this corresponds to an arbitrary linear rescaling of the radial variable r.

This choice of scaling also corresponds to the arbitrary scaling in the definition of the first Lyapunov coefficient of the Hopf bifurcation, an ambiguity that is removed by choosing a convention (Kuznetsov, 2010). The quantity that remains invariant under a linear rescaling, and matters for dynamical behavior such as synchronization, is the ratio a/b.

There is more than one valid way to consider this scaling choice. We could choose a scaling of r that preserves the same numerical value for the oscillation amplitude as the original system. That would require defining "amplitude" in the original system. As explained in Pikovsky et al. (2003) there are multiple valid ways to do that. This may result in different scaling methodologies depending on the problem at hand. The benefit of this approach is that more physical properties will be quantitatively preserved by the transformation.

Alternatively we can make an unambiguous choice that works for all systems, and rescale r so that the coefficient of the r^3 term in the normal form is always -1. This is the approach we have taken below, by applying the change of variables,

$$R = \sqrt{a}\, r. \tag{3.34}$$

3.4.2 Coupling Hopf normal form systems

When a simplified oscillator model $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ is used to study synchronization in networks, with a coupled network of N simple oscillators,

$$\dot{\mathbf{x}}_j \;=\; \mathbf{f}(\mathbf{x}_j) + \sum_{i=1}^{N} c_{ij}\mathbf{h}(\mathbf{x}_i, \mathbf{x}_j), \tag{3.35}$$

the coupling function $\mathbf{h}$ between the simple oscillators is either chosen heuristically or inferred from a more detailed biophysical model. In Section 3.5 below we demonstrate that close to a bifurcation where oscillation dynamics are fundamentally 2-dimensional, finding the physically appropriate parameters and form of coupling is necessary for correctly capturing the synchronization behavior of a system, showing the value of the latter approach.

For a 1-dimensional phase oscillator model such as a network of Kuramoto oscillators $\dot{\theta} = \omega$, the most commonly used coupling function is

$$h(\theta_i, \theta_j) = \sin(\theta_i - \theta_j). \tag{3.36}$$

For phase oscillators this is a reasonable heuristic choice in many cases. If the natural frequencies ω_i, ω_j are close, with weak, symmetric coupling $h(\theta_i, \theta_j) = h(\theta_j, \theta_i)$, then averaging around the cycle

to remove dependence on absolute phase yields an odd function of the phase difference $h(\theta_i, \theta_j) = q(\theta_i - \theta_j) = -q(\theta_j - \theta_i)$ so that the first term of its Fourier series is $c\sin(\theta_i - \theta_j)$ (Pikovsky et al., 2003). However higher order terms can be important (Hansel et al., 1993).

Phase oscillator models with coupling function derived explicitly by reduction of a biophysical model include Schuster and Wagner (1990), Brown et al. (2004), Daffertshofer and van Wijk (2011), and Ashwin and Rodrigues (2016).

For 2-dimensional reduced phase-amplitude oscillators such as the supercritical Hopf normal form $\dot{z} = (\gamma\epsilon + \omega i)z + (a + bi)z|z|^2$ or 5th order generalized Hopf (Bautin) normal form $\dot{z} = (\gamma\epsilon + \omega i)z + (a + bi)z|z|^2 + (c + di)z|z|^4$, several heuristic coupling schemes are often chosen in the literature:

Linear diffusive coupling: $h(z_i, z_j) = c.(z_i - z_j)$.
Diffusive coupling acts to bring the states of coupled systems closer together (Pikovsky et al., 2003; Aronson et al., 1990). Named *diffusive* as the earliest application was to chemical oscillators literally coupled by diffusion (e.g. Torre, 1974). This choice is often made when the system is known to exhibit phase and amplitude synchronization and it is desired to capture this in the simplified model. By allowing the scalar coefficient c to be complex a phase offset can be included in the interaction. Examples of normal form models in neuroscience using this coupling include Benjamin et al. (2012) and Terry et al. (2012).

Direct coupling: $h(z_i, z_j) = z_i$.
This form, where coupling depends linearly on the source system alone, is named *direct* following Aronson et al. (1990). In neuroscience models this coupling form reflects the fact that axonal propagation of action potentials between neuron populations is unidirectional. Examples from computational neuroscience include Koppert et al. (2012) and Kalitzin et al. (2014)

Periodic direct coupling: $h(z_i, z_j) = \mathrm{Re}(z_i)$.
If the output of a source region is a scalar quantity, such as a population mean firing rate, periodic coupling reflects the fact that the physical coupling mechanism may directly depend on that quantity, but not its rate of change. An example of its use in computational neuroscience is Koppert et al. (2014).

Despite the fact that axonal coupling is unidirectional, in the nervous system the resulting mean post-synaptic potential of a target population receiving input is to some extent dependent on the state of the target cells. At the single cell level this is due to both synaptic and dendritic processes including facilitation (leading to greater response to coupling input), refractory states of ion channels, and spike rate adaptation (both leading to reduced response to coupling input). At the population level, local recurrent inhibitory and excitatory connections can also regulate excitation in the target region, mediating the population response to coupling input.

A biological justification of linear diffusive coupling between oscillations, or any other coupling that involves both source and target regions, rests on these processes. Now these biophysical features can be captured in neural mass models (for example spike rate adaptation and recurrent inhibition are implemented in Moran et al. (2007)) but whether the effect is approximately equivalent to linear diffusive coupling of phase and amplitude is not obvious.

The transformation method presented below gives a way to approach this question. We derive a first approximation to the coupling function between two normal form systems by transforming the coupling terms from the original biological model to the 2-dimensional normal form coordinates.

3.5 Application to neural mass model and synchronization

We compare a network of coupled 8-dimensional Jansen-Rit neural mass models with two different simplified networks of 2-dimensional Hopf normal forms, showing that synchronization behavior of the original neural mass network is not preserved in the simplified model if standard parameters are assigned heuristically from the topological normal form with diffusive coupling, but synchronization behavior is preserved if the parameters and coupling are constrained by transformation.

As a neural mass model of a single cortical region capable of oscillation the Jansen-Rit model (Jansen et al., 1993; Jansen and Rit, 1995) was introduced in Section 2.2.1. When treated as a deterministic system of first order ordinary differential equations in the variables $\mathbf{v} = (v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8)^T$, with arbitrary firing rate inputs $p(t)$ and $u(t)$ to pyramidal and spiny stellate populations respectively,

this model is

$$\dot{\mathbf{v}} \;=\; \mathbf{f}(\mathbf{v},t) \;=\; \begin{bmatrix} v_5 \\ v_6 \\ v_7 \\ v_8 \\ H_e\kappa_e(\gamma_1 S(v_2-v_3)+u(t)) - \kappa_e^2 v_1 - 2\kappa_e v_5 \\ H_e\kappa_e(\gamma_2 S(v_1)+p(t)) - \kappa_e^2 v_2 - 2\kappa_e v_6 \\ H_i\kappa_e\gamma_4 S(v_4) - \kappa_i^2 v_3 - 2\kappa_i v_7 \\ H_e\kappa_e\gamma_3 S(v_2-v_3) - \kappa_e^2 v_4 - 2\kappa_e v_8 \end{bmatrix}, \tag{3.37}$$

$$\text{where} \quad S(v) = \frac{2e_0}{1+\exp[\rho_1(\rho_2-v)]}. \tag{3.38}$$

As shown by Grimbert and Faugeras (2006) and reproduced by numerical continuation in Section 2.3.1, for the case $u(t) = 0$ and standard parameter values listed in Table 2.1, a supercritical Hopf bifurcation occurs as the pyramidal input $p(t)$ is increased beyond $p_c \approx 89.8\ \mathrm{s}^{-1}$ (labeled point H1 in Figure 2.2).

In fact for this system it is possible to solve for the location of the equilibrium point $\mathbf{v}_c$ and Hopf bifurcation point p_c to arbitrary precision, directly from Eq. (3.37). The working for this is shown in Appendix A.1. In the case of Hopf point H1 the result is:

$$p_c = 89.829\,\mathrm{s}^{-1}, \quad \mathbf{v}_c = (13.208, 20.165, 13.425, 3.3021, 0, 0, 0, 0)^{\mathrm{T}}\ \mathrm{mV}. \tag{3.39}$$

3.5.1 Network of coupled Jansen-Rit systems

Here we construct an all-to-all coupled network of five neighboring cortical regions operating just beyond the onset of oscillations (input firing rate $p_c + \epsilon = 104.3\,\mathrm{s}^{-1}$ so that $\gamma_\epsilon \epsilon = 0.2$), each modeled with a separate instance of the 8-dimensional Jansen-Rit model above.

To model input to the pyramidal population from neighboring cortical regions, similarly to Goodfellow et al. (2011) we use the same nonlinear synaptic coupling between populations that is used between

populations in the Jansen-Rit model itself, so the total pyramidal input firing rate $p(t)$ to region j is

$$p_j(t) = p_0(t) + \sum_{i \neq j} c_{ij}\gamma_1 S(v_{i2}(t) - v_{i3}(t)), \tag{3.40}$$

where c_{ij} is the (directed) connectivity strength from region i to region j, $v_{i2} - v_{i3}$ is the mean pyramidal membrane potential of the ith source region, with an extrinsic white noise process $p_0(t) = \langle p_0 \rangle + \sigma_0 \xi(t)$ representing input from other brain regions which are not being explicitly modeled in the network. No extra delay (beyond the delay implied by the characteristic synaptic response time) is introduced in the coupling as we are modeling directly adjacent cortical regions. (cf. Jansen and Rit (1995) which models distal regions.)

At the bifurcation point, each region in the network produces a total output firing rate $\phi_c = \gamma_1 S(v_{c2} - v_{c3})$ which for point H1 is $\phi_c \approx 406\,\mathrm{s}^{-1}$. When simulating a single Jansen-Rit model the source of input that brings the system to the threshold p_c is not modeled dynamically, instead it is modeled as a static mean value $\langle p \rangle$ of a Gaussian white noise process, as in Chapter 2. But now with a network model some of the input to each region is being modeled explicitly by the other regions, so for the subsystems simultaneously to be near the onset of oscillations $p(t) = p_c + \epsilon$ requires the remaining extrinsic input that is not modeled dynamically, $\langle p_0 \rangle$, to be reduced accordingly, i.e. $\langle p_0 \rangle = p_c + \epsilon - \sum_{i \neq j} c_{ij}\phi_c$. For the present simulation beyond the bifurcation, that gives $\langle p_0 \rangle = 103.8\,\mathrm{s}^{-1}$. The standard deviation σ_0 of the extrinsic input process is reduced in proportion to its mean, giving $\sigma_0 = 0.420\,\mathrm{s}^{-1}$.

For later use, we rewrite this network model with the coupling from other regions as a separate term instead of writing this input as part of $p(t)$. Then the equation for the jth region is

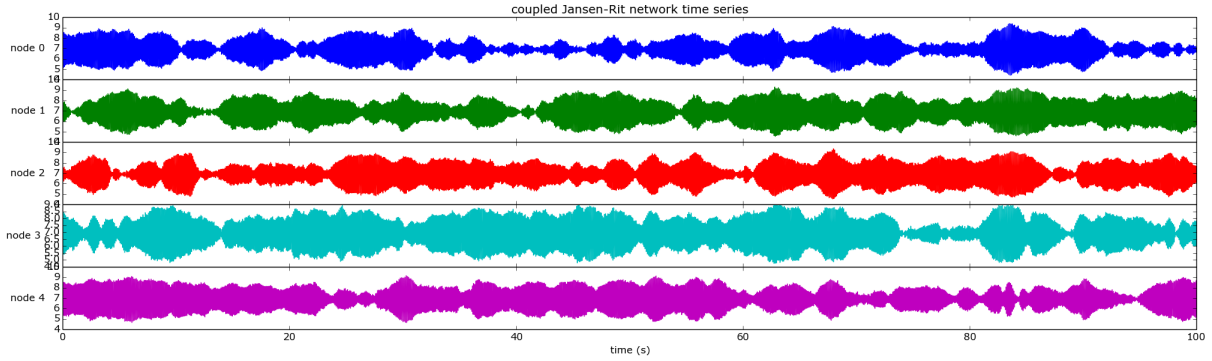
$$\dot{\mathbf{v}}_j \;=\; \mathbf{f}(\mathbf{v}_j) + \sum_{i \neq j} c_{ij}\mathbf{h}(\mathbf{v}_i, \mathbf{v}_j), \tag{3.41}$$

where $\mathbf{f}(\mathbf{v}_j)$ is the function $\mathbf{f}(\mathbf{v}, t)$ of Eq. (3.37) but now setting $u(t) = 0, p(t) = p_0$ (bifurcation point H1), and with coupling function

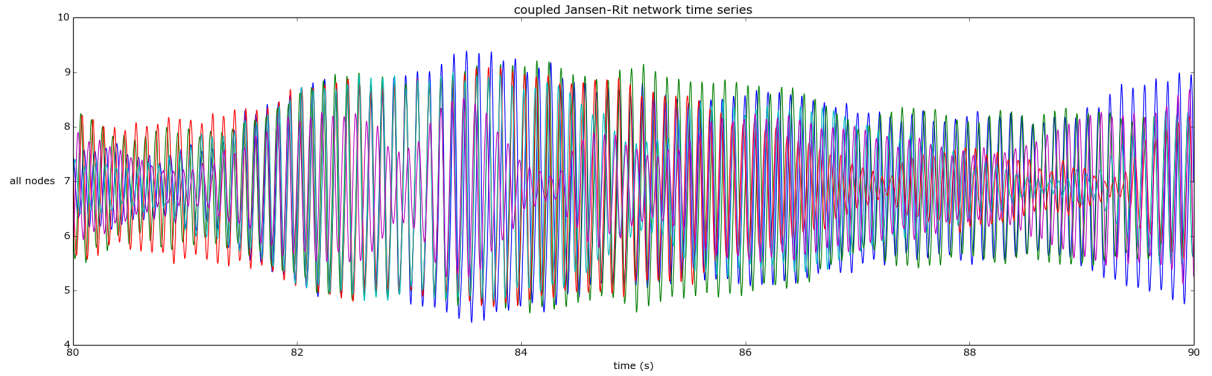
$$\mathbf{h}(\mathbf{v}_i, \mathbf{v}_j) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ H_e \kappa_e \gamma_1 S(v_{i2} - v_{i3}) \\ 0 \\ 0 \end{bmatrix}. \tag{3.42}$$

A network of $N = 5$ coupled Jansen-Rit models was simulated with all-to-all bidirectional coupling: $c_{ij} = \frac{0.09}{N}$ for $i \neq j, c_{ii} = 0, p_0 = 16.23\,\mathrm{s}^{-1}$.

A sample of the resulting output for the pyramidal cell populations at each node in the network is shown in Figure 3.3. The dominant behavior is alpha-frequency oscillations in each region, with oscillation amplitude that waxes and wanes on a slower time scale of seconds. The oscillation amplitude is correlated between nodes in the network. The regions tend to synchronize in phase when mean oscillation amplitude is large, then desynchronize when mean amplitude is small.



A Amplitude relationship between the five nodes, showing full simulation time range $t = 0\,\mathrm{s}$ to $t = 100\,\mathrm{s}$.



B Phase relationship between the five nodes, showing the short time interval $t = 80\,\mathrm{s}$ to $t = 90\,\mathrm{s}$.

Figure 3.3: A network of 5 coupled Jansen-Rit systems, plotting the pyramidal output time series at each node. This shows dynamic changes of synchronization of both amplitude and phase between nodes.

3.5.2 Network of Hopf normal forms, standard parameters

Having identified the bifurcation type (supercritical Hopf) and base frequency of oscillations (from the imaginary eigenvalues at bifurcation) we might attempt qualitatively to capture this network synchronization behavior in a simpler model, by constructing a network of simple 2-dimensional Hopf normal form systems of the same frequency:

$$\begin{aligned} \dot{r} &= \gamma_\epsilon \epsilon r + a r^3 \\ \dot{\theta} &= \omega + \omega_\epsilon \epsilon + b r^2, \end{aligned} \tag{3.43}$$

or in complex form, setting $z = re^{i\theta}$ and $\gamma = \gamma_\epsilon + \mathrm{i}\,\omega_\epsilon$,

$$\dot{z} = (\gamma\epsilon + \omega i) z + (a + bi) z |z|^2, \tag{3.44}$$

where angular frequency $\omega = 65.2\,\mathrm{s}^{-1}$. Without a method to constrain the other parameters of the supercritical Hopf normal form, we use the standard values of its topological normal form: $\gamma = 1$, $a = -1$ and $b = 0$. To facilitate synchronization of both phase and amplitude, we then couple these normal form systems with a linear diffusive coupling function proportional to the difference of the complex amplitudes $z_i - z_j$, so for the jth region, the simple network model is given by

$$\dot{z_j} = (\epsilon + \omega i)z_j - z_j|z_j|^2 + \sum_{i \neq j} c_{ij}(z_i - z_j) + \eta\sigma_0\xi(t). \tag{3.45}$$

To make the extrinsic noise comparable to the Jansen-Rit network simulation, the same noise process $\sigma_0\xi(t)$ is used, then multiplied by $\eta = 0.60 - 0.019\,\mathrm{i}$ to linearly transform it to the plane of oscillation.

We simulate an all-to-all coupled network of $N = 5$ adjacent regions beyond the onset of oscillations, with $\epsilon = 0.2$ and $c_{ij} = \frac{0.09}{N}$ for $i \neq j, c_{ii} = 0$, the same network structure as before.

3.5.3 Network of normal forms, parameterized by transformation

To contrast with the approach of Section 3.5.2 above, we create a third network of simple 2-dimensional Hopf normal form systems, this time with its parameters and coupling constrained by Poincaré normal form transformation of the Jansen-Rit model, instead of assigning standard parameters from the topological normal form.

Computing the ϵ-dependent inner-product normal form from the Jansen-Rit model Eq. (3.37) near Hopf point H1, the working at each step is presented in Appendix A.2. After a linear rescaling to fix $a = -1$, the resulting normal form is:

$$\dot{z} = (\gamma\epsilon + \omega i)z + (a + bi)z|z|^2, \tag{3.46}$$

where $\omega = -65.2, \quad a = -1, \quad b = 5.71,$ and $\gamma = 0.0138 - 0.0831i.$

From the nonlinear coupling term $\mathbf{h}(\mathbf{v}_i, \mathbf{v}_j)$ of the original Jansen-Rit model, Eq. (3.42), we take a bilinear approximation of the coupling vector field near Hopf point H1,

$$\mathbf{h}(\mathbf{v}_i, \mathbf{v}_j) \approx \mathbf{h}(\mathbf{v}_c, \mathbf{v}_c) + (\frac{\partial}{\partial \mathbf{v}_i}\mathbf{h}(\mathbf{v}_i, \mathbf{v}_j)).(\mathbf{v}_i - \mathbf{v}_c) + (\frac{\partial}{\partial \mathbf{v}_j}\mathbf{h}(\mathbf{v}_i, \mathbf{v}_j)).(\mathbf{v}_j - \mathbf{v}_c), \tag{3.47}$$

then transform this to the 8-dimensional normal form coordinate system, so that the first two components drive the oscillations of Eq. (3.46). The details at each step are shown in Appendix A.3. The resulting network model with first order coupling is:

$$\dot{z_j} = (\gamma\epsilon + \omega i)z_j + (a + bi)z_j|z_j|^2 + \sum_{i \neq j} c_{ij}(c_0 + c_1\text{Re}(z_i) + c_2\text{Im}(z_i)) + \eta\sigma_0\xi(t), \tag{3.48}$$

where $c_0 = 244 - 7.77\text{i}, c_1 = 87.9 - 2.80\text{i}$ and $c_2 = -38.7 + 1.23\text{i}$.

In contrast to the heuristic choice of coupling function of Section 3.5.2, in this system the coupling does not depend on the difference between source and target regions, but on the source region alone. This arises naturally from transforming the synaptic coupling of the Jansen-Rit model, which has the same property, i.e. $\frac{\partial}{\partial \mathbf{v}_j}\mathbf{h}(\mathbf{v}_i, \mathbf{v}_j) = 0$. Note also that in general a constant term c_0 appears in the coupling.

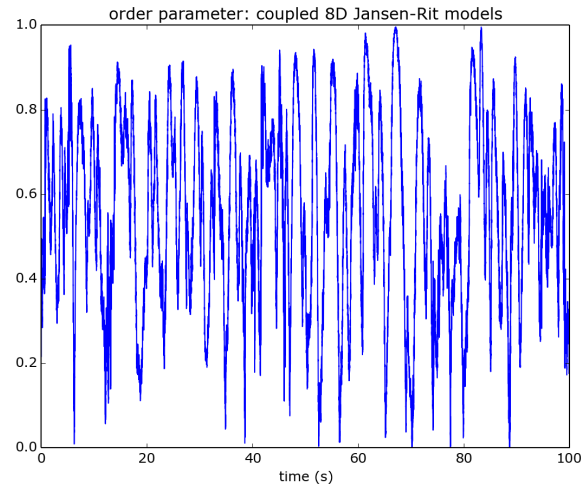
Here the extrinsic noise $\eta\sigma_0\xi(t)$ is the same as in the previous simulation, with $\eta = 0.60 - 0.19\,\text{i}$ as before. Again we simulate an all-to-all coupled network of $N = 5$ adjacent regions beyond the onset of oscillations, with $\gamma_\epsilon\epsilon = 0.2$ and $c_{ij} = \frac{0.09}{N}$ for $i \neq j, c_{ii} = 0$, as with the previous simulations.

3.5.4 Results

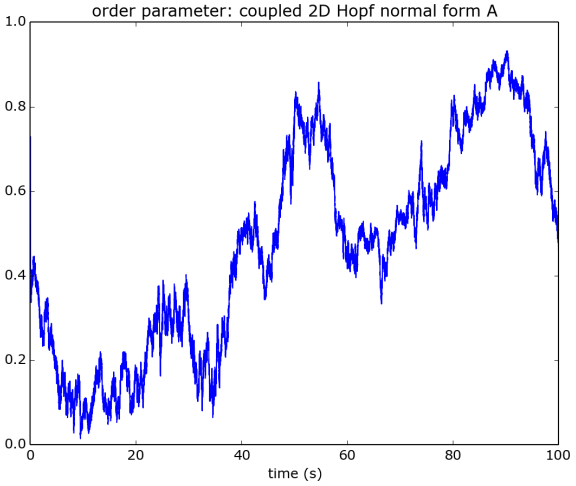
Comparing the full 40-dimensional network of Jansen-Rit neural masses and the two Hopf normal form networks, all three models give oscillations that are similar when viewed individually, that is alpha frequency oscillations modulated by a slowly varying amplitude. But the network synchronization of the three models is not the same. To view the time course of synchronization for each network, we plot at each point in time the modulus of the phase order parameter,

$$\rho(t) = \frac{1}{N}\sum_{j=1}^{N} e^{\text{i}\theta_j(t)}, \tag{3.49}$$

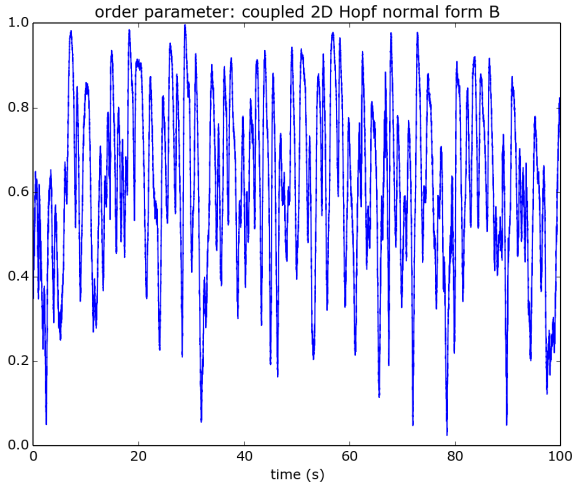
which lies between 0 (for a uniform phase distribution) and 1 (for an ensemble of identical phase). The results for synchronization of the three different all-to-all coupled networks with the same coupling strength and network structure are shown in Figure 3.4:



A Network of 5 coupled Jansen-Rit systems



B Network of 5 coupled Hopf normal form systems with zero shear and $c.(z_2 - z_1)$ coupling



C Network of 5 coupled Hopf normal form systems parameterized by transformation

Figure 3.4: Order parameter showing the degree of synchronization for (A) a network of five 8-dimensional Jansen-Rit models near Hopf bifurcation point H1 with synaptic coupling, (B) network of 2-dimensional Hopf normal forms with parameters of the topological normal form and $c.(z_2 - z_1)$ coupling and (C) network of 2-dimensional Hopf normal forms with parameters and coupling derived by transformation.

Repeating the simulations with all three networks implemented as ordinary differential equations without any noise (setting $\sigma_0 = 0$) gave similar results.

3.6 Discussion

Using a Poincaré normal form transformation to parameterize the simplified model and its coupling allowed the collective synchronization behavior to be qualitatively preserved in a network simulation, whereas assigning standard parameters from the topological normal form and heuristic choice of coupling did not.

This approach can be applied also to parameterize reduced models near other local bifurcations of equilibria, such as systems containing saddle-node, Bautin or cusp bifurcations.

To make analytical predictions about network synchronization from the normal form models, a further reduction from the normal form to a one-dimensional phase oscillator model is sometimes possible. This is viable in cases where the coupling (and all other perturbations such as noise) are weak enough (Ashwin and Rodrigues, 2016), or else the limit cycle far enough beyond bifurcation, that the amplitude dynamics are negligible (Kuramoto and Haken, 1984). We have seen that close to a supercritical Hopf bifurcation the dynamics are fundamentally two-dimensional. How far from the bifurcation point this remains true depends on the bifurcation parameter scale γ_ϵ, one of the parameters computed above in the Poincaré normal form.

Other normal form algorithms exist beyond the simplest one presented here. Those methods are more computationally efficient, and in the case of Lie-theory based algorithms give both the forward and inverse transformation explicitly (Murdock, 2003). Having the inverse transformation then for example allows a phase response curve of the normal form to be transformed back to the original biological coordinates (Brown et al., 2004).

The reduced amplitude stability that occurs near a Hopf bifurcation has two separate effects on oscillator synchronization. First is the dependence of phase coupling between oscillators on their amplitudes, as shown by Daffertshofer and van Wijk (2011) and evident here. Second is the dependence within a single oscillator of oscillator frequency variance on its amplitude variance, a relationship determined by the normal form shear parameter b. In systems with large b, the amplitude variance near bifurcation causes greater frequency variance which in turn reduces synchrony near bifurcation, for the same coupling strength. Whereas in a system with small b the first effect, but not the second, occurs.

This demonstrates why selecting normal form parameters, such as shear b, bifurcation parameter scale γ_ϵ and coupling function h approximately to match the system being modeled is important for studies of network synchrony: the collective behavior depends upon them, even though as isolated oscillators their flows are topologically equivalent. For simulations using normal form models, this paper has shown one way to constrain parameters of the simple model by transformation of a detailed biophysical model.

Chapter 4

Transformation of stochastic systems explains the effect of noise near bifurcations

For many physical and biological systems that can transition between oscillation and no oscillation, such as healthy and pathological brain rhythms, nanomechanical resonators and financial cycles in the economy, the most common generic dynamical mechanism underlying the onset of oscillations is a Hopf bifurcation. We demonstrate a new method to analyze any dynamical model close to a Hopf bifurcation, to derive how random noise perturbations close to the transition point will manifest as modified phase and amplitude statistics of the oscillations.

In contrast to mere simulation, this analytical approach derives the precise functional form for how phase and amplitude noise terms depend upon parameters of the original model. This facilitates making predictions from models about how oscillation statistics will change as physical or biological parameters are changed in experiment.

To do this, normal form transformation expresses the system in a new coordinate system where the emerging oscillations (in the absence of noise) are symmetrical circles in a flat plane. By using stochastic calculus to also transform the noise perturbations to this same coordinate system, their first order effects on oscillation phase and amplitude are made explicit.

In this paper we first give a brief overview of the method and transform a simple three-dimensional example system before examining the general case in Section 4.4, including a criterion for the reduced

system to be a good approximation. Section 4.5 examines the meaning of each term in the resulting standard simplified form. In Chapter 5 we demonstrate that not only phase and amplitude probability distributions but also several temporal properties are preserved by the transformation. Finally Chapter 6 applies this method to the Jansen-Rit neural mass model to explain the phenomena reported in Aburn et al. (2012).

4.1 Background

For computational neuroscience gradually to arrive at better models, it is necessary to compare model predictions with experimental data.

But in modeling complex systems such as the brain, usually not every dynamical variable of a model can be observed directly in experiment. Sometimes what can be observed is a projection of the system state onto a lower-dimensional subspace. A good example of this is a neural mass model, which aims to use 6 to 10 dynamical variables to capture the collective dynamics of a small region of cortical tissue, but where what is measurable in a clinical or experimental setting is often a single variable (such as a one-dimensional time series for each brain region derived from EEG, MEG or local field potential recording). Oscillation statistics from time series, such as the rate of phase diffusion or the distribution of oscillation amplitude and its mean reversion time, may be measurable in experiment, while the evolution of all dynamic variables may not be.

When such a system is at a stable equilibrium, its local behavior is approximately linear, and in analyzing experimental time series from a stable equilibrium the theory of linear time-invariant systems applies. In this case any perturbations from the equilibrium (including noise driven perturbations) are quickly dissipated and decay exponentially in time back to the stable state. Thus only response behaviors on short time scales are observable. Because response amplitude is damped the response is also more easily masked by measurement noise (such as the activity of other brain regions). In this stable regime the best connection that can be made between the model and observed time series is to determine from the model a transfer function relating frequencies in the input perturbation to frequencies in the resulting system output (e.g. Robinson et al., 2002; Robinson et al., 2004).

Similarly when the system is in a regime of stable oscillations on a limit cycle, perturbations away from this attractor also decay exponentially back to the limit cycle. If a one-dimensional time series can be observed in experiment, the dominant behavior is the oscillation at the base frequency. Beyond this the relative strength of harmonics (integer multiples of the oscillator base frequency) in the output may distinguish between different classes of nonlinear oscillator (Chatterjee, 2009).

But it is close to the bifurcation between these two regimes, where linear stability is temporarily lost, that the susceptibility of the system to intrinsic or extrinsic perturbations is greatest. As the response is not so quickly dissipated the real system is thus likely to divulge more information about the underlying dynamics. Here behaviors on longer time scales can potentially be observed. Where models explain a transition of behavior in the real system, and it is possible experimentally to observe the system through this transition, this offers the best opportunity to measure the statistical behavior of the real system and compare to models.

Rather than predicting behavior of the system at a single point in parameter space only (which can be done by simulating the model using those specific parameters multiple times and collecting statistics from the repeated simulations) our method goes further by allowing to predict analytically from the model the functional form of how the phase and amplitude noise terms of leading order depend upon biological parameters.

4.2 Method overview: pipeline of transformations

Near a Hopf bifurcation, the original system in general will have non-circular oscillations that emerge on an arbitrarily curved two-dimensional surface embedded within the N-dimensional state space. This makes it difficult to determine exactly how noise will affect oscillation phase and amplitude. But after applying the pipeline of transformations, in the new coordinates the system will have symmetrical circular oscillations on a flat surface. The leading order noise effects will be expressed directly in terms of phase and amplitude.

To achieve this, the original system is transformed to a simpler system using the steps summarized in Figure 4.1.

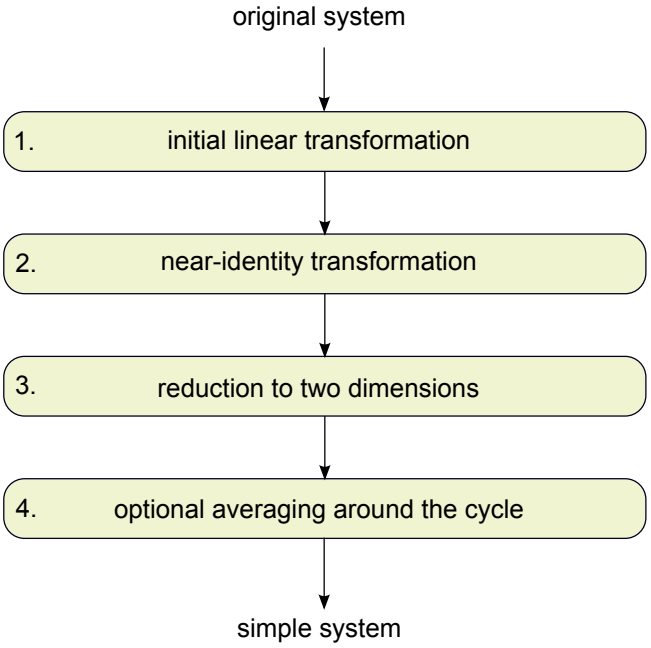


Figure 4.1: Pipeline of transformations

The meaning of each of these steps is as follows:

1. translation and initial linear transformation.

Translate the bifurcating equilibrium point to the origin and apply a linear transformation to align the coordinate system with the (generalized) eigenspaces. After this step the Jacobian of the system is in Jordan real form.

2. near-identity transformation

Approximate the deterministic system near the origin with a Taylor series. Then find the normal form transformation described in Section 3.3.2 which maps the deterministic system to a simple normal form system with symmetrical circular oscillations in a flat plane. Having found the transformation that makes the deterministic oscillations simple, now apply the same transformation to the full stochastic system including noise perturbations, using Stratonovich calculus.

3. reduction to two dimensions

For a deterministic system, the first two steps will de-couple the center eigenspace variables from the stable variables. By contrast, when noise is present in general they will still be coupled through stochastic terms that depend on the stable variables. As shown in Section 4.4.3 these stochastic coupling terms scale as $\frac{\sigma^2}{\lambda^{1.5}}$, so that the system de-couples to a good approximation

when the noise intensity is small and/or the stable eigenspaces very stable. In this step these stochastic coupling terms are either omitted or replaced with a constant intensity noise term with the same root-mean-square noise intensity. The result is a 2-dimensional normal form system with additive and multiplicative noise terms.

4. optional: averaging around the cycle

If the behavior of interest is of a slower time scale than the oscillation period then short-term dependency on the phase can be removed. To do this the reduced system is expressed as a Fokker-Planck equation and averaged around the cycle.

(Alternatively in systems where this separation of time scales does not hold, so that the phase dependent dynamics remain important, rather than averaging the phase dependence is instead simplified by taking leading terms of a Fourier series expansion.)

4.3 Example with two noise perturbations

We first demonstrate the method with a simple 3-dimensional example system. This is based on the deterministic system of Section 3.2, but now with two independent Gaussian white noise perturbations:

$$\begin{aligned} \dot{x}_1 &= \epsilon x_1 - 9x_2 - x_1x_3 + \sigma_1\xi_1(t) \\ \dot{x}_2 &= 9x_1 + \epsilon x_2 - x_2x_3 \\ \dot{x}_3 &= -2x_3 + 4(x_1^2 + x_2^2) + \sigma_3\xi_3(t). \end{aligned} \tag{4.1}$$

Because the noise is additive, that is noise intensity independent of the system state $\mathbf{x}$, its expression above using Langevin symbols $\xi_1(t), \xi_3(t)$ is unambiguous. This example will be useful to illustrate how additive noise in the original n-dimensional system gives rise to both additive and multiplicative noise acting on the 2-dimensional phase and amplitude dynamics of oscillation.

In the steps below we will change to a new coordinate system by applying continuously invertible transformations to the system. Automating these calculations is simplest using Stratonovich stochastic calculus, so a preparatory step is to express the system in Stratonovich form. Because the noise here

is additive this is trivial for this first example:

$$\begin{aligned} dx_1 &= (\epsilon x_1 - 9x_2 - x_1 x_3)dt + \sigma_1 \circ dW_1 \\ dx_2 &= (9x_1 + \epsilon x_2 - x_2 x_3)dt \\ dx_3 &= (-2x_3 + 4(x_1^2 + x_2^2))dt + \sigma_3 \circ dW_3. \end{aligned} \tag{4.2}$$

We shall contrast the effect on this system of the two different noise perturbations in Eq. (4.2). Each has constant intensity and constant direction in phase space: $\sigma_1 \circ dW_1$ is in a tangential direction aligned with the x_1 axis and $\sigma_3 \circ dW_3$ is in the axial direction aligned with the x_3 axis, as shown in Figure 4.2. Note that in this example the noise perturbation does not have circular symmetry — we are applying noise in the direction of x_1 but not x_2. From Figure 4.2, defining an oscillation phase $\theta = \arg(x_1 + \mathrm{i}\, x_2)$ we can see that the noise $\sigma_1 \circ dW_1$ (red, aligned with the x_1 direction) will perturb both phase and amplitude of oscillations during each cycle as the phase θ evolves around the circle. It will affect oscillation amplitude most when $\sin\theta$ is small, and will affect oscillation phase most when $\cos\theta$ is small.

We can also see from this figure that the axial noise $\sigma_3 \circ dW_3$ (pink, aligned with the x_3 direction) will not directly perturb the phase θ. (It indirectly affects the phase evolution, by perturbing the amplitude,

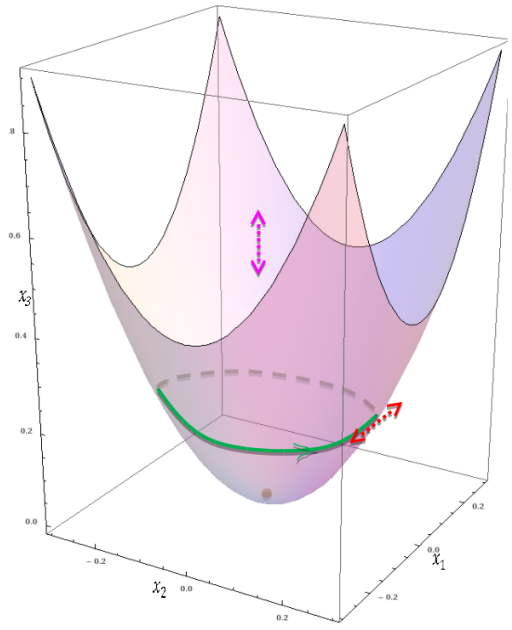


Figure 4.2: Example system showing the shape of the center manifold surface and direction of two noise perturbations $\sigma_1 \circ dW_1$ (red, in the plane of oscillation) and $\sigma_3 \circ dW_3$ (pink, in the axial direction).

which in turn changes the susceptibility of the phase to tangential noise.)

As illustrated in Figure 4.3, because the center manifold is curved, the effect of the axial noise perturbation $\sigma_3 \circ dW_3$ on the oscillation amplitude depends on the state $\mathbf{x}$ of the system. This noise has little effect on oscillation amplitude when the amplitude is small, but a greater effect when amplitude is larger, due to different alignment with the center manifold surface. (More precisely we shall see this depends on the size of $\frac{\partial R}{\partial x_3}$ where R is the radial variable in a normal form coordinate system which parameterizes the curved surface.) This shows in graphical form how simple additive noise in the original system becomes multiplicative (state dependent) noise when restricted to the two-dimensional phase-amplitude dynamics on the center manifold surface.

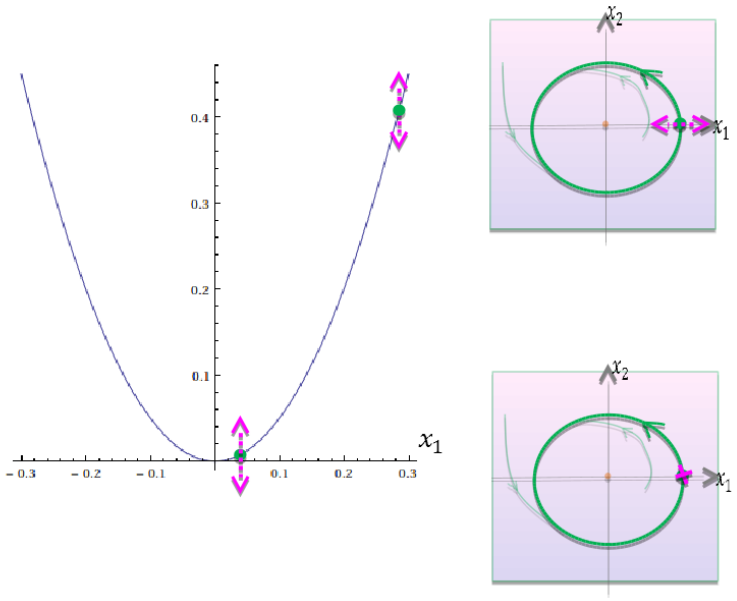


Figure 4.3: Due to curvature of the center manifold, the effect of the axial noise $\sigma_3 \circ dW_3$ depends on the state of the system. It has little effect on amplitude dynamics when amplitude is small (bottom) but a greater effect when amplitude is larger (top) due to alignment with the center manifold surface.

We next demonstrate that transforming the stochastic system to Poincaré normal form coordinates explicitly provides the same information. Whereas in this simple 3D example we can see what will happen directly from Figures 4.2 and 4.3, in a real world application, which is typically asymmetrical and N-dimensional, the method proposed here is useful to derive explicitly the effect of noise on phase and amplitude dynamics.

Steps 1 and 2 - normal form transformation

To find the transformation which maps the unperturbed oscillations to circles in a flat plane we set noise intensities σ_1 and σ_3 to zero and examine the corresponding deterministic system. It is the same system that was treated in Section 3.3. The normal form algorithm is applied (as detailed in Section 3.2), to find a smooth change of variables which maps the deterministic system Eq. (3.3) to its normal form. The result is the near-identity transformation to new variables y_1, y_2, y_3,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 + y_1 y_3/2 + y_1 y_3^2/8 \\ y_2 + y_2 y_3/2 + y_2 y_3^2/8 \\ y_3 + 2y_1^2 + 2y_2^2 \end{bmatrix}. \tag{4.3}$$

Having found this transformation, we now apply it to transform the full stochastic system Eq. (4.2), truncating series to third order in $|(y_i, \sigma_j, \sqrt{\epsilon})|$ to give

$$\begin{aligned} dy_1 &= (\epsilon y_1 - 9y_2 - 2y_1^3 - 2y_1 y_2^2)dt + \sigma_1(1 + 2y_1^2 - y_3/2 + y_3^2/8) \circ dW_1 \\ &\qquad - \sigma_3 \frac{y_1}{2} \circ dW_3 + O(|\mathbf{y}|^4) \\ dy_2 &= (9y_1 + \epsilon y_2 - 2y_1^2 y_2 - 2y_2^3)dt + 2y_1 y_2 \sigma_1 \circ dW_1 - \sigma_3 \frac{y_2}{2} \circ dW_3 + O(|\mathbf{y}|^4) \\ dy_3 &= (-2y_3 + 4y_1^2 y_3 + 4y_2^2 y_3)dt + \sigma_1(-4y_1 - 8y_1^3 - 8y_1 y_2^2 + 2y_1 y_3 - y_1 y_3^2/2) \circ dW_1 \\ &\qquad + \sigma_3(1 + 2y_1^2 + 2y_2^2) \circ dW_3 + O(|\mathbf{y}|^4) \end{aligned} \tag{4.4}$$

in Stratonovich form. This step will be discussed in Sections 4.4.1 and 4.4.2.

Expressing Eq. (4.4) in polar coordinates,

$$\begin{aligned} dr &= (\epsilon r - 2r^3)dt + \sigma_1(1 + 2r^2 - y_3/2 + y_3^2/8)\cos\theta \circ dW_1 - \frac{r}{2}\sigma_3 \circ dW_3 + O(r^4) \\ d\theta &= 9\, dt + \sigma_1(-\frac{1}{r} + \frac{y_3}{2r} - \frac{y_3^2}{8r})\sin\theta \circ dW_1 + O(r^3) \\ dy_3 &= (-2y_3 + 4r^2 y_3)dt + \sigma_1(-4r + 2ry_3)\cos\theta \circ dW_1 + \sigma_3(1 + 2r^2) \circ dW_3 + O(r^4). \end{aligned} \tag{4.5}$$

Step 3 - reduction to two dimensions

Note that for the stochastic system, the evolution of the oscillating variables r and θ is still dependent on the stable variable y_3, through stochastic coupling terms such as $\sigma_1(y_3/2)\cos\theta \circ dW_1$ and $\sigma_1(y_3^2/8)\cos\theta \circ dW_1$. We now estimate the size of these terms, to make clear in what situations these y_3 terms are small compared to the other terms affecting the oscillations. The stable variable y_3 to lowest order follows an Ornstein-Uhlenbeck process,

$$dy_3 = -2y_3 dt + \sigma_3 \circ dW_3, \tag{4.6}$$

which tends to a stationary Gaussian distribution of mean zero and standard deviation $\frac{\sigma_3}{\sqrt{2}\sqrt{|-2|}} = \sigma_3/2$. This standard deviation gives an estimate of the root-mean-square average size of y_3, i.e. how close the system stays to the deterministic center manifold. Thus the coupling term $-\sigma_1(y_3/2)\cos\theta \circ dW_1$ scales as $\sigma_1\sigma_3/4$, quadratic in the noise intensities.

In Section 4.4.3 we shall prove that these leading order coupling terms scale as $\frac{\sigma^2}{\lambda^{1.5}}$ in general, where $-\lambda$ is the real part of the weakest stable eigenvalue, so that the coupling terms will be small if either the noise is weak or the stable manifold very stable.

The contribution of the coupling term $-y_3/2$ is an order σ^2 zero-mean fluctuation of the noise intensity about the positive value $1 + 2r^2$, we approximate $1 + 2r^2 - y_3/2 \approx 1 + 2r^2$ provided $\sigma_1\sigma_3/4 \ll 1$.

The higher order coupling terms such as $y_3^2/8$ have non-zero mean. But they scale as σ^3/λ^3, so are negligible in our asymptotic approximation.

The approximation above is invalidated if either the stable eigenvalue is very weak, so that mean-reversion of the y_3 process is slow compared to the oscillation period, or if there is resonance between

the eigenvalue of the Hopf bifurcation and a stable eigenvalue. Neither caveat applies to this example system.

Thus the first two equations of Eq. (4.5) decouple to a good approximation when $\sigma_1\sigma_3/4 \ll 1$ and we can examine the oscillations as the two-dimensional system in the (r, θ) plane,

$$\begin{aligned} dr &= (\epsilon r - 2r^3)dt + \sigma_1(1+2r^2)\cos\theta \circ dW_1 - \frac{r}{2}\sigma_3 \circ dW_3 \\ d\theta &= 9\, dt - \sigma_1 \frac{1}{r}\sin\theta \circ dW_1. \end{aligned} \tag{4.7}$$

Step 4 - optional averaging around the cycle

Provided the timescale of oscillations (here of order ~ 1) is short compared with the characteristic timescales $\sim \frac{1}{\epsilon}, \frac{1}{\sigma_1^2}, \frac{1}{\sigma_3^2\epsilon^2}$ of the radial dynamics, Eq. (4.7) can be averaged around the cycle. This is done by changing to the rotating frame $\phi = \theta - 9t$ and considering the corresponding Fokker-Planck equation,

$$\begin{aligned} \partial_t\, p(r,\phi,t) = [&- \partial_r((\epsilon + \frac{\sigma_1^2}{2}(3+\cos(2(9t+\phi))) + \frac{\sigma_3^2}{8})r - 2r^3 + \frac{\sigma_1^2}{4r}(1-\cos(2(9t+\phi))) \\ &+ \frac{1}{2}\partial_r^2(\sigma_1^2(1+2r^2)^2\cos^2(9t+\phi) + \frac{\sigma_3^2 r^2}{4}) - \partial_r\partial_\phi\frac{1+2r^2}{2r}\sigma_1^2\sin(2(9t+\phi)) \\ &+ \frac{1}{2}\partial_\phi^2\frac{\sigma_1^2\sin^2(9t+\phi)}{r^2}]\, p(r,\phi,t). \end{aligned} \tag{4.8}$$

Averaging the Fokker-Planck operator around the cycle gives:

$$\partial_t p(r,\phi,t) \approx \left[-\partial_r((\epsilon + \frac{3\sigma_1^2}{2} + \frac{\sigma_3^2}{8})r - 2r^3 + \frac{\sigma_1^2}{4r}) + \frac{1}{2}\partial_r^2(\frac{\sigma_1^2}{2}(1+2r^2)^2 + \frac{\sigma_3^2}{4}r^2) + \frac{1}{2}\partial_\phi^2\frac{\sigma_1^2}{2r^2}\right] p(r,\phi,t), \tag{4.9}$$

for which one choice of corresponding Stratonovich SDE is

$$\begin{aligned} dr &= ((\epsilon + \frac{\sigma_1^2}{2})r - 2r^3 + \frac{\sigma_1^2}{4r})dt + (\frac{\sigma_1}{\sqrt{2}} + \frac{8\sigma_1^2+\sigma_3^2}{4\sqrt{2}\sigma_1}r^2) \circ dW_A \\ d\phi &= \frac{\sigma_1}{r\sqrt{2}} \circ dW_B, \end{aligned} \tag{4.10}$$

chosen by Cholesky decomposition of the Fokker-Planck diffusion matrix. This step will be discussed in Section 4.4.4. Returning to the non-rotating frame and linearly rescaling the radial variable as

$R = \sqrt{2}r$ we arrive at the standard form:

$$dR = ((\epsilon + \frac{\sigma_1^2}{2})R - R^3 + \frac{\sigma_1^2}{2R})dt + (\sigma_1 + \frac{8\sigma_1^2 + \sigma_3^2}{8\sigma_1}R^2) \circ dW_A$$
$$d\theta = 9\, dt + \frac{\sigma_1}{R} \circ dW_B. \tag{4.11}$$

In the standard form it can be seen that for this particular system:

- Only the tangential noise (σ_1) directly perturbs the phase, resulting in phase diffusion.
- The axial noise (σ_3) contributes to multiplicative noise affecting the amplitude R.

This matches the picture given in Figure 4.3. The Mathematica package accompanying this paper automates the above steps to transform and approximate an arbitrary Stratonovich system at the point of Hopf bifurcation.

4.4 Applying transformation to noise-perturbed system: general method

Starting with the dynamical system and noise written as an Ito stochastic differential equation,

$$\mathbf{dx} = \mathbf{f}(\mathbf{x}, \epsilon)dt + \sigma G(\mathbf{x})\mathbf{dW}, \tag{4.12}$$

here G in general is a $d \times p$ matrix, so that $\sigma\, G\, \mathbf{dW}$ is a d-dimensional vector noise term driven by p independent Weiner noise processes. Thus the off-diagonal entries of G determine whether the noise driving different components of the equation is correlated or uncorrelated.

To transform this system we first consider the corresponding deterministic system, setting noise intensity σ to zero.

In **step 1** we translate the equilibrium point to the origin and apply an initial linear transformation to put the Jacobian of the system into Jordan real form.

During this step, if the Jacobian of the system at bifurcation is ill-conditioned with respect to eigenvalue computation (that is if the ratio of largest to smallest singular values is large), we can optionally linearly transform to balance the matrix first, using the method of Parlett and Reinsch (1969). This is particularly important if applying the normal form algorithm of Murdock (2003) chapter 2.1, in which a projection operator is computed directly from the spectrum of the Jacobian matrix, in cases where the matrix is far from orthogonal.

In **step 2** we apply the normal form algorithm using computer algebra as described in Chapter 3, to find explicitly a smooth invertible transformation,

$$\mathbf{x} = \mathbf{y} + \mathbf{q}(\mathbf{y}), \tag{4.13}$$

which maps the deterministic system at the bifurcation point, $\mathbf{dx} = \mathbf{f}(\mathbf{x}, 0)\, dt$, to a normal form system with symmetrical circular oscillations in a flat plane, $\mathbf{dy} = \mathbf{g}(\mathbf{y})\, dt$.

4.4.1 Choice of asymptotic approximation of stochastic vector fields

The right-hand-side of the system (4.12) is a random contravariant vector field representing a stochastic flow on the state space. Throughout the steps below we work with truncated power series approximations to the coefficient functions of the field at each step.

There are three different kinds of small quantities in the field expressions which we are approximating: the state space variables x_i and r, the noise intensities σ_i and the bifurcation parameter ϵ. Hence deterministic terms and the noise coefficient functions are approximated as multivariate power series in *all* of these small quantities, truncating the series according to a chosen asymptotic scaling between the small quantities.

In general there is flexibility in choosing asymptotic approximations. The choice does not simply involve a total ordering of possible approximations by accuracy. Instead approximation must be seen as a tradeoff between accuracy and the simplicity/tractability of the resulting system. The relative accuracy of two approximations cannot be inferred from the asymptotic order of terms alone without fixing the values of series coefficients. Our goal therefore is to produce asymptotic approximations that are as simple as possible while retaining the main noise induced effects to lowest order.

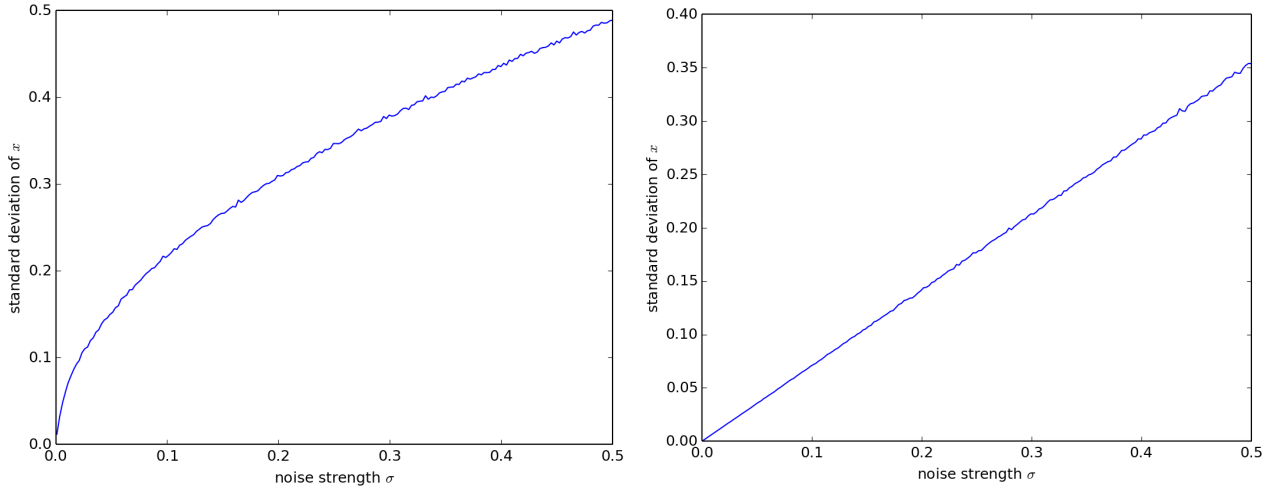
Sometimes in the literature a low-noise asymptotic scaling relationship is introduced without comment which makes the mathematics tractable. But in a stochastic system we are not free to choose an asymptotic scaling arbitrarily. As noise intensities σ_i are increased, the average values of the state space variables x_i will be pushed out, with a definite relationship. That is the physical system takes a specific path towards the origin in the asymptotic limit, and other choices of scaling are unphysical.

For series approximations when approaching a supercritical Hopf bifurcation, to choose an asymptotic scaling in a principled way consider the two scalar processes,

$$dx = -x^3\,dt + dW, \quad \text{and} \tag{4.14}$$

$$dx = -x\,dt + dW. \tag{4.15}$$

Very close to a supercritical Hopf bifurcation the real part of the bifurcating eigenvalues is negligible compared to cubic terms and the weaker cubic stability regime of Eq. (4.14) applies to the amplitude of oscillations. Here oscillation amplitude increases with the square root of noise intensity, $O(\sigma) = O(r^2)$, as shown in Figure 4.4 below. But while still approaching bifurcation the linear term is significant and the linear stability of the Ornstein-Uhlenbeck regime of Eq. (4.15) applies, with oscillation amplitude scaling linearly with noise intensity: $O(\sigma) = O(r)$. In our system the true scal-



A Scaling of $dx = -x^3\,dt + dW$ B Scaling of $dx = -x\,dt + dW$

Figure 4.4: Asymptotic scaling of oscillation amplitude with noise intensity, for cubic regime (A) and Ornstein-Uhlenbeck regime (B).

ing behavior smoothly transitions between these two regimes as the bifurcation is approached. Thus to approximate accurately the noise-induced behavior through the bifurcation we employ the stricter

of the above scalings, $O(\sigma) = O(r)$. As will be seen, this retains several noise-induced terms that would be omitted if assuming $O(\sigma) = O(r^2)$ throughout. As this is a supercritical Hopf bifurcation, for $\epsilon > 0$ the oscillation amplitude also scales with the square root of the bifurcation parameter ϵ.

For the reasons above a default asymptotic scaling $O(\sigma_j^2) = O(x_i^2) = O(\epsilon)$ is used when truncating series in the computer algebra computations below. Our accompanying Mathematica package implements operations between multivariate power series truncated according to user-defined asymptotic scalings, so the above default scaling can be changed if appropriate for specific systems (or local bifurcations other than Hopf).

4.4.2 Transformation of noise coefficient matrix G

Having found the desired transformation $\mathbf{x} = \mathbf{y} + \mathbf{q}(\mathbf{y})$, which renders the deterministic oscillations simple, we next use this new coordinate system to view the perturbed system (4.12), including all noise terms and bifurcation parameter.

When applying the transformation to the stochastic terms Stratonovich calculus will be used so that the matrix of noise coefficients G will transform algebraically in the same way as the deterministic field $\mathbf{f}$. Therefore if Eq. (4.12) contains any multiplicative noise (that is, G depends on $\mathbf{x}$) then it must first be expressed in Stratonovich form,

$$\mathbf{dx} = \underline{\mathbf{f}}(\mathbf{x}, \epsilon)dt + \sigma G(\mathbf{x}) \circ \mathbf{dW}, \tag{4.16}$$

$$\text{where} \qquad \underline{f}_i = f_i - \frac{1}{2}\sum_{j=1}^{p}\sum_{k=1}^{d} G_{kj}\frac{\partial G_{ij}}{\partial x_k}. \tag{4.17}$$

Applying the near-identity change of coordinates Eq. (4.13) to transform the coefficients $\underline{\mathbf{f}}$ and G of Eq. (4.16) as contravariant fields,

$$\begin{aligned}
& \mathbf{dx} = \mathbf{dy} + (D\mathbf{q}(\mathbf{y}))\mathbf{dy}, \\
\Longrightarrow \qquad & \mathbf{dy} = (I + D\mathbf{q}(\mathbf{y}))^{-1}\mathbf{dx}, \\
\Longrightarrow \qquad & \mathbf{dy} = \quad (I + D\mathbf{q}(\mathbf{y}))^{-1}\underline{\mathbf{f}}(\mathbf{y} + \mathbf{q}(\mathbf{y}), \epsilon)dt \\
& \qquad\qquad + \sigma(I + D\mathbf{q}(\mathbf{y}))^{-1}G(\mathbf{y} + \mathbf{q}(\mathbf{y})) \circ \mathbf{dW}.
\end{aligned} \tag{4.18}$$

This uses the fact that $(I+D\mathbf{q}(\mathbf{y}))$ is invertible in a neighborhood of the origin since the transformation is near-identity, that is, $\mathbf{q}(\mathbf{y}) = O(|\mathbf{y}|^2)$.

Multivariate series approximations of the coefficient functions of the deterministic and stochastic parts of Eq. (4.18) are computed symbolically to mth order in $|(\mathbf{y}, \sigma, \sqrt{\epsilon})|$ from mth order series approximations of $\mathbf{q}(\mathbf{y})$, $\underline{\mathbf{f}}(\mathbf{x})$ and $G(\mathbf{x})$.

It can be seen from Eq. (4.18) that even if the original system had only additive noise (that is G independent of $\mathbf{x}$), after applying the transformation any curvature in the center manifold surface will now manifest as multiplicative noise in the transformed system (4.18). This can be proved as follows. Since the normal form transformation maps the original center manifold to a flat plane, it is exactly when the original center manifold is curved that the normal form transformation Eq. (4.13) has nonlinear terms involving the center variables y_1 and y_2. This implies that the factor $(I + D\mathbf{q}(\mathbf{y}))^{-1}$ in Eq. (4.18) will be a function of $\mathbf{y}$ that involves y_1 and y_2 so that additive noise terms in the original system give rise to new multiplicative noise terms in the transformed system that depend upon the center variables y_1 and y_2.

Since the normal form transformation is equal to the identity I to first order, the transformed system will also retain the existing additive noise terms, unchanged by the near-identity transformation.

4.4.3 Criterion for approximate decoupling

For the unperturbed deterministic system, the transformation causes the two oscillating variables y_1 and y_2 to be decoupled from the rest of the system so their evolution does not depend on the stable variables $y_3, y_4, \ldots$. But note that in the stochastic system this is not the case. Consider the case where the noise in the original system is additive, so that G is a constant matrix. In the transformed system, besides additive noise terms $\sigma G \circ \mathbf{dW}$, the equations for y_1 and y_2 will have multiplicative noise terms $\sigma((I + D\mathbf{q}(\mathbf{y}))^{-1} - I)G \circ \mathbf{dW}$ and the first two components of this vector field may still depend upon the stable variables $y_3, y_4, \cdots$. Under what circumstances is this stochastic coupling to the stable system negligible, so that the oscillations can be analyzed as a separate two-dimensional stochastic system? For the case where the noise in the original system is additive, we will show that any stochastic coupling terms involving the stable y_i scale as $\frac{\sigma^2}{\lambda^{1.5}}$ where σ is the noise strength and

$-\lambda$ is the real part of the weakest stable eigenvalue of A. Thus the system approximately decouples provided either the noise is small or the stable eigenspace is sufficiently stable.

The following steps justify this criterion for decoupling. First we estimate the distribution of the stable variables $y_3, y_4, \cdots, y_d$. The stable y_i to lowest order follow a multivariate Ornstein-Uhlenbeck process,

$$dy_i = \sum_{j=3}^{d} A_{ij} y_j dt + \sigma \sum_{j=1}^{p} G_{ij} dW_j \qquad i = 3..d. \tag{4.19}$$

This tends to an asymmetrical stationary probability distribution in which each $y_3..y_d$ has a smaller variance $\langle y_i^2 \rangle$ than those of the *symmetrical* Ornstein-Uhlenbeck process,

$$dy_i = -\lambda y_i dt + \sigma\gamma dW \qquad i = 3..d, \tag{4.20}$$

where $-\lambda$ is the real part of the weakest stable eigenvalue of A and $\gamma = \max_i \sqrt{\sum_{j=1}^{p} G_{ij}^2}$.
But the symmetrical process Eq. (4.20) tends to a multivariate Gaussian distribution where each y_i has mean zero and standard deviation $\frac{\sigma\gamma}{\sqrt{2\lambda}}$. This standard deviation therefore gives an upper bound for the root-mean-square average size of the stable y_i in Eq. (4.19), showing that the y_i scale as $\frac{\sigma}{\lambda^{0.5}}$ or smaller, in the limit of small σ and large λ. This then gives a conservative estimate that the stochastic coupling monomial terms $\sigma a y_1^{m_1} y_2^{m_2} y_3^{m_3} ... \circ dW$ scale as $\frac{\sigma^2}{\lambda^{0.5}}$ or smaller. The stronger estimate of $\frac{\sigma^2}{\lambda^{1.5}}$ is now achieved by showing that the scalar coefficient a itself scales as $\frac{1}{\lambda}$.

The multiplicative noise terms $\sigma((I + D\mathbf{q}(\mathbf{y}))^{-1} - I)G \circ \mathbf{dW}$ have the property that any monomial $\prod_k y_k^{m_k}$ appearing in the first two components of the field is always multiplied by a scalar coefficient a that scales as $(\sum_k m_k \lambda_k)^{-1}$. To prove this it is sufficient to show that the transformation $\mathbf{q}(\mathbf{y})$ has this property.

As discussed in Section 3.3.2 $\mathbf{q}(\mathbf{y})$ satisfies the homological equation,

$$\mathfrak{L}_B \, \mathbf{q}(\mathbf{y}) = \mathbf{f}(\mathbf{y}) - \mathbf{g}(\mathbf{y}), \tag{4.21}$$

where $B = A$ in the semisimple case and $B = A^\dagger$ in the non-semisimple case. The essence of the proof below is that the Jordan block structure of B induces a corresponding block structure on the operator

$\mathfrak{L}_B$, whence Eq. (4.21) restricted to the subspace of each block then gives the required property for $\mathbf{q}(\mathbf{y})$.

Let $\mathbf{b}_\alpha = \mathbf{e}_i \prod_j y_j^{m_j}$ and $\mathbf{b}_\beta = \mathbf{e}_k \prod_l y_l^{m_l}$ be two basis vectors in the standard basis $\{\mathbf{b}_\gamma\}$ of $\mathcal{V}$, as introduced in Section 3.3.2. Denote by $\mathbf{d}_\alpha$ the dual space basis vector corresponding to $\mathbf{b}_\alpha$, that is $\mathbf{d}_\alpha = \mathbf{e}^i \prod_j (\frac{1}{m_j!} \frac{\partial^{m_j}}{\partial y_j^{m_j}})$, writing an inner product $(\mathbf{d}_\alpha, \mathbf{b}_\beta) = (\mathbf{d}_\alpha \cdot \mathbf{b}_\beta)|_{\mathbf{y}=\mathbf{0}} = \delta_{\alpha\beta}$. Then the matrix $\hat{L}$ representing $\mathfrak{L}_B$ with respect to this basis is

$$\begin{aligned}\hat{L}_{\alpha\beta} &= (\mathbf{d}_\alpha, \mathfrak{L}_B \mathbf{b}_\beta) = (\mathbf{d}_\alpha, \mathfrak{L}_B(\mathbf{e}_k \prod_l y_l^{m_l})) \\ &= (\mathbf{d}_\alpha, \sum_{p,r} \mathbf{e}_p B_{pr} \delta_{rk} \prod_l y_l^{m_l} - \mathbf{e}_k \sum_{h,r} \frac{\partial}{\partial y_h} (\prod_l y_l^{m_l}) B_{hr} y_r) \qquad \text{by definition of } \mathfrak{L}_B. \end{aligned} \tag{4.22}$$

As B is in Jordan real form, the real parts of its stable eigenvalues ($-\lambda_i$ for $i \geq 3$) appear only on its diagonal: $B_{ii} = -\lambda_i$. We can now rearrange Eq. (4.22) to separate the effects of the diagonal and non-diagonal parts of B:

$$\begin{aligned}\hat{L}_{\alpha\beta} &= (\mathbf{d}_\alpha, B_{kk}\mathbf{e}_k \prod_l y_l^{m_l} + \sum_{p\neq k} B_{pk}\mathbf{e}_p \prod_l y_l^{m_l} - \mathbf{e}_k \sum_h m_h (\prod_l y_l^{m_l}) y_h^{-1} (B_{hh} y_h + \sum_{r\neq h} B_{hr} y_r)) \\ &= (\mathbf{d}_\alpha, B_{kk}\mathbf{b}_\beta + \sum_{p\neq k} B_{pk}\mathbf{e}_p \prod_l y_l^{m_l} - \sum_h B_{hh} m_h \mathbf{b}_\beta - \sum_h \sum_{r\neq h} m_h B_{hr} \mathbf{e}_k (\prod_l y_l^{m_l}) y_h^{-1} y_r) \\ &= \delta_{\alpha\beta}(B_{kk} - \sum_h B_{hh} m_h) + \sum_{p\neq k} B_{pk} (\mathbf{d}_\alpha, \mathbf{e}_p \prod_l y_l^{m_l}) - \sum_h \sum_{r\neq h} m_h B_{hr} (\mathbf{d}_\alpha, \mathbf{e}_k (\prod_l y_l^{m_l}) y_h^{-1} y_r). \end{aligned} \tag{4.23}$$

Eq. (4.23) shows that only diagonal elements $\hat{L}_{\alpha\alpha}$ depend on the λ_i, and further that the Jordan block structure of B imposes a corresponding structure on the operator $\mathfrak{L}_B$ as follows: The third term in Eq. (4.23) shows that $\mathfrak{L}_B$ can map one basis vector $\mathbf{b}_\beta$ to another $\mathbf{b}_\alpha$ by replacing a y_h with a y_r but only when $B_{hr} \neq 0$ so that y_h and y_r belong to the same Jordan block of B, implying that $B_{hh} = B_{rr}$. From the first term of Eq. (4.23) this then implies that for these two basis vectors $\hat{L}_{\alpha\alpha} = \hat{L}_{\beta\beta}$. Now for any basis vector $\mathbf{b}_\eta = \mathbf{e}_i \prod_j y_j^{m_j}$ define the subspace $\mathcal{W}_\eta$ as the span of all basis vectors $\mathbf{b}_\zeta = \mathbf{e}_k \prod_l y_l^{m_l}$ with $k = i$ and $\hat{L}_{\zeta\zeta} = \hat{L}_{\eta\eta}$. Then the whole space $\mathcal{V}$ is the direct sum of the $\mathcal{W}_\eta$ and $\mathfrak{L}_B$ is a vector space endomorphism on each of the $\mathcal{W}_\eta$. Let P_η denote the projection operator into the subspace $\mathcal{W}_\eta$. Since $\mathfrak{L}_B \mathcal{W}_\eta \subseteq \mathcal{W}_\eta$, $\mathfrak{L}_B$ and P_η commute.

Now suppose that $\mathbf{q}(\mathbf{y})$ contains a term $a\mathbf{b}_\beta = a\mathbf{e}_k \prod_l y_l^{m_l}$ where $k \in \{1,2\}$, $m_l > 0$ for some $l \geq 3$ and real coefficient $a \neq 0$. Near a Hopf bifurcation the first block of A is $\left[\begin{smallmatrix} \epsilon & -\omega \\ \omega & \epsilon \end{smallmatrix}\right]$, so for $k \in \{1,2\}$ Eq. (4.23) gives

$$\hat{L}_{\beta\beta} = \epsilon(1 - m_1 - m_2) + \sum_{h=3}^{n} m_h \lambda_h. \tag{4.24}$$

Projecting the homological equation Eq. (4.21) into the subspace $\mathcal{W}_\beta$,

$$\begin{aligned} \mathfrak{L}_B P_\beta \mathbf{q}(\mathbf{y}) &= P_\beta \mathfrak{L}_B\, \mathbf{q}(\mathbf{y}) \\ &= P_\beta \mathbf{f}(\mathbf{y}) - P_\beta \mathbf{g}(\mathbf{y}), \end{aligned} \tag{4.25}$$

where $\mathbf{f}(\mathbf{y})$ are the nonlinear deterministic terms in the original system and $\mathbf{g}(\mathbf{y})$ are the terms in the transformed system. But $P_\beta \mathbf{g}(\mathbf{y}) = \mathbf{0}$, since the deterministic system decouples so $\mathbf{g}(\mathbf{y})$ does not contain terms involving y_l for any $l \geq 3$ in its first two components. Note also that $\mathbf{f}(\mathbf{y})$ is independent of the λ_i.

Now express $P_\beta \mathbf{q}(\mathbf{y})$ and $P_\beta \mathbf{f}(\mathbf{y})$ in terms of those standard basis elements $\{\mathbf{b}_\gamma\}_{\gamma=1}^{\nu}$ that span $\mathcal{W}_\beta$ thus: $P_\beta \mathbf{q}(\mathbf{y}) = \sum_\gamma a_\gamma \mathbf{b}_\gamma$ and $P_\beta \mathbf{f}(\mathbf{y}) = \sum_\gamma c_\gamma \mathbf{b}_\gamma$, with $\mathbf{b}_1 = \mathbf{b}_\beta$ and $a_1 = a$. Eq. (4.25) then implies:

$$\begin{aligned} & \sum_{\eta=1}^{\nu} \hat{L}_{\zeta\eta} a_\eta = c_\zeta \quad \text{for } \zeta = 1..\nu, \\ \implies \qquad & \Lambda a_\zeta + \sum_{\eta \neq \zeta} \hat{L}_{\zeta\eta} a_\eta = c_\zeta, \end{aligned} \tag{4.26}$$

where

$$\begin{aligned} \Lambda = \hat{L}_{\zeta\zeta} &= \hat{L}_{\beta\beta} \\ &= \epsilon(1 - m_1 - m_2) + \sum_{h=3}^{n} m_h \lambda_h. \end{aligned} \tag{4.27}$$

Since the a_η satisfy the linear system Eq. (4.26) and the c_ζ and off-diagonal entries of $\hat{L}$ are independent of Λ, this implies that the a_η are rational functions of Λ. Therefore for each ζ, $a_\zeta = O(\Lambda^s)$ for some $s \in \mathbb{Z}$ as $\Lambda \to \infty$. We now prove that $a_\zeta = O(\Lambda^{-1})$ for all $\zeta = 1..\nu$.

For contradiction, assume there is some ζ for which $\lim_{\Lambda\to\infty} |\Lambda a_\zeta| = \infty$.

But for any integer $s \geq -1$, if $\exists\zeta : \lim_{\Lambda\to\infty} |\frac{a_\zeta}{\Lambda^s}| = \infty$ then Eq. (4.26) implies that

$$\begin{aligned}\lim_{\Lambda\to\infty} |\sum_{\eta\neq\zeta} \frac{\hat{L}_{\zeta\eta} a_\eta}{\Lambda^{s+1}}| &= \lim_{\Lambda\to\infty} |\frac{c_\zeta - \Lambda a_\zeta}{\Lambda^{s+1}}| \\ &= \lim_{\Lambda\to\infty} |\frac{a_\zeta}{\Lambda^s}| \\ &= \infty. \end{aligned} \tag{4.28}$$

This implies that $\exists\eta : \lim_{\Lambda\to\infty} |\frac{a_\eta}{\Lambda^{s+1}}| = \infty$ (because the $\hat{L}_{\zeta\eta}$ are independent of Λ). By induction on s it follows that for all integers $s \geq -1$ there exists an η for which $\lim_{\Lambda\to\infty} \frac{a_\eta}{\Lambda^s} = \infty$. This contradicts the fact that there is a finite set of a_η, each of which is of integer order in Λ. This completes the proof that $a_\zeta = O(\Lambda^{-1})$ for all ζ, and in particular $a = a_1 = O(\Lambda^{-1}) = O(\lambda^{-1})$ where λ is the weakest of the stable λ_i. Thus the coefficient a scales as $\frac{1}{\lambda}$ and so combined with the scaling $\frac{\sigma}{\lambda^{0.5}}$ of the stable y_i discussed above, any stochastic coupling terms $\sigma a y_1^{m_1} y_2^{m_2} y_3^{m_3} ... \circ dW$ in the transformed system scale as $\frac{\sigma^2}{\lambda^{1.5}}$ or smaller.

4.4.4 Averaging

This is the final step in the pipeline. After reduction to two dimensions, the system still retains full detail of the dependence of the noise effects on the phase θ.

Now if the period of oscillations is short compared to the time scales of the behaviors of interest (such as phase diffusion or amplitude reversion behavior), then the reduced system can be averaged around the cycle to give a further simplified system with circular symmetry (independent of the phase θ) that still accurately reflects all behaviors that occur on a longer time scale. This provides a weak model valid on longer time scales with dynamics independent of phase.

To implement the averaging we first transform the system to a frame that is rotating at the base linear frequency ω of oscillations at the origin, introducing the relative phase variable $\phi \equiv \theta - \omega t$. In this frame both amplitude and relative phase distributions evolve slowly compared to the period.

Next we move to the weak approximation, deriving from the SDE the corresponding Fokker-Planck equation. This loses information about individual sample paths but retains the same drift and diffusion of phase and amplitude. Where the Ito or Stratonovich noise coefficient matrix G retained terms to order 3, we retain terms to order 6 in the Fokker-Planck diffusion matrix $D = GG^T$. In Fokker-Planck form we can now apply the Krylov-Bogolyubov method of averaging (Bogoliubov and Mitropolsky, 1961). For this to be a valid approximation depends upon the separation of time scales described above.

After averaging, we can express the process as an SDE again, for comparison with the original system. For dimension $d \geq 2$, to a single diffusion process expressed by a Fokker-Planck equation there corresponds an infinite set of SDEs, which are stochastically equivalent in the weak sense. These all satisfy $GG^T = D$ where G is the noise coefficient matrix of the SDE and D is the diffusion matrix of the FPE (Gardiner, 2010). We propose two different ways to do this:

1. Diagonalize the diffusion matrix D of the Fokker-Planck equation (if possible), then take the square root of the diagonal entries to obtain G. Then $GG^T = G^2 = D$.

2. Compute the Cholesky decomposition of D as $D = GG^T$.

Method 2 has two benefits: unlike method 1 it will always be applicable, since the diffusion matrix D is positive semi-definite but may not be diagonalizable, and because the Cholesky decomposition results in a matrix G which is upper triangular, in general the resulting SDE is therefore simpler, having less non-zero terms than that from method 1.

The form of each term is also simpler for method 2, as it avoids the terms with square roots within square roots that are produced by method 1. But this is not a major concern as we take a power series asymptotic approximation and retain just the lowest order multiplicative noise terms in both cases.

For both methods, in general the simplified SDE system will have two driving Wiener processes.

Note that while choices of SDE from a FPE may be stochastically equivalent in that their probability distributions are identical at all times, they may yet differ in the temporal properties of sample paths. The fidelity of the averaged weak approximation to the temporal properties of the original system is one question examined in Chapter 5.

Alternative to averaging: Fourier expansion

To handle those cases where the short-time angular dependency of the noise effects is important to the oscillation dynamics (for example a system where oscillations are slow so the time scale separation does not hold) then can take the leading one or two terms of a Fourier series expansion around the cycle, instead of averaging, as an alternative way to simplify the angular dependency of the noise effects. This will give a system that is not quite as simple but retains better fidelity for the behavior at short time scales. This case has not been implemented in our code.

4.5 Interpretation of terms in the simplified system

The end result of this pipeline is a two-dimensional Stratonovich system in a standard form:

$$\begin{aligned} dr &= (-r^3 + (\gamma_\epsilon \epsilon + \gamma_\sigma \sigma^2) r + h\sigma^2 \frac{1}{r})\, dt \quad + \quad (p + qr^2)\sigma \circ dW_A \\ d\theta &= (\omega + \omega_\epsilon \epsilon + \omega_\sigma \sigma^2 + br^2)\, dt \quad + \quad s_a r \sigma \circ dW_A \quad + \quad (\frac{p}{r} + s_b r)\sigma \circ dW_B. \end{aligned} \tag{4.29}$$

Comparing this to the deterministic normal form,

$$\begin{aligned} dr &= (-r^3 + \gamma_\epsilon \epsilon r)\, dt \\ d\theta &= (\omega + \omega_\epsilon \epsilon + br^2)\, dt \end{aligned} \tag{4.30}$$

we see that the new parameters γ_σ, ω_σ and h describe how the noise modifies the deterministic dynamics. $\gamma_\sigma \sigma^2$ is a shift of the bifurcation point caused by the noise. $\omega_\sigma \sigma^2$ is a shift of the oscillation frequency caused by the noise. The term $h\sigma^2 \frac{1}{r}$ affects the small end of the amplitude distribution, shifting the modal amplitude away from the origin, an effect that is most significant before the deterministic bifurcation point.

The remaining parameters p, q, s_a and s_b describe the additive and multiplicative noise in phase and amplitude. Here p quantifies the additive noise, while q, s_a and s_b give the multiplicative noise effects to lowest order.

4.6 Discussion

With this method we can now view the system in two forms: In the original form, the connection between equation terms and their biological meaning is clear. Whereas in the transformed system, the connection between equation terms and oscillator properties is clear. In this way the transformation makes explicit the effect of a noise perturbation in the original equations on phase and amplitude dynamics near the bifurcation, something that is not obvious from the original equations.

When one or more biological parameters of the original equations are left as variables or symbols rather than numerical values during the transformation process, we get explicit expressions for how each term in the standard form depends on the biological parameters. This allows making predictions from the model about the change in oscillation statistics to be expected as a biological parameter is modified. This is a new way to test models against experimental data.

Chapter 5

Properties preserved by transformation of stochastic systems

5.1 Overview

When a system is viewed in the simplified form, in which oscillations are circular and lie in a flat plane, each term has a known effect on phase or amplitude dynamics as described in Section 4.5. By retaining one or more biological parameters from the original system as symbols with unspecified values when performing the transformation to the simplified system, the result is explicit expressions for the functional dependence of each simple oscillation term on biological parameters.

Thus the transformed system gives direct insight into how biological parameters affect a range of properties of the stochastic phase and amplitude dynamics, provided that the properties of interest are preserved by the transformation.

But which properties are preserved through the transformation?

In this chapter we consider two broad types of statistics of the oscillating system and use large scale simulation of sample paths to examine whether they are preserved by each stage of the transformation pipeline:

Probability distributions of phase and amplitude

For both phase and amplitude we will examine both the limiting stationary distribution as time $t \to \infty$ and also the initial transient evolution of the distributions in time. Specifically we examine phase diffusion (the increase of the spread of the phase distribution over time) and the probability distribution of the oscillation amplitude.

Temporal properties

Even if two stochastic processes are equivalent in the weak sense, that is their probability distributions are the same at all times, the behavior of individual sample paths in time may differ between the two processes. To characterize the temporal behavior of sample paths we look at the distribution of oscillation periods, the autocorrelation and power spectra of the amplitude, and the distribution of mean reversion times.

The next sections describe in detail these six statistical properties to be examined. These properties were all selected in advance before performing the simulations, excepting power spectral density which was added subsequently to provide a clearer view of the information present in the autocorrelation function.

5.1.1 Defining phase and amplitude in the original system

After the normal form transformation (step 2) the oscillations are approximately circular and lie in the plane of the first two dynamical variables y_1 and y_2. So from this stage onwards in our transformation pipeline it is straightforward to define phase and amplitude by projection to this plane as the argument and modulus of $y_1 + \mathrm{i}\, y_2$ for example.

But for the original system it is not trivial to define what is meant by phase and amplitude of the oscillations. As discussed in Pikovsky et al. (2003) there are various equally valid approaches that could be used, giving non-equivalent definitions of phase and amplitude.

We aim to measure the statistics of oscillation dynamics in a way that is consistent and comparable all the way along the pipeline of transformations. Therefore to define the instantaneous phase and amplitude in all cases we use a common definition. We take the projection of the state into the one-dimensional subspace corresponding to y_1 (a vector in the plane of the oscillations in the normal form

system), subtract any linear trend then define the phase and amplitude to be the argument and modulus of the analytical signal computed by the Hilbert transform, as introduced in Section 1.2.1.

By a direction "corresponding to" y_1 we mean that the unit dual vector for the y_1 direction in the normal form system is transformed covariantly back to the original coordinates of the biophysical model, to find the dual vector of the equivalent direction in the original system.

This choice of definition for phase and amplitude by the Hilbert transform of a projection has the benefit of matching common practice for analysis of electrophysiological time series, as discussed in Section 1.2.1. However it gives a small artifact for the first few time steps due to boundary conditions. This may be improved by padding with an odd (reflected) signal at $t = 0$ though we did not implement this for these simulations.

5.1.2 Phase and amplitude probability distributions

Probability distribution of phase: the time course of phase diffusion

When the system starts from a known state, the distribution of the oscillation phase is initially a Dirac delta distribution on the circle $-\pi < \theta \leq \pi$. As time progresses the distribution is at first close to a wrapped normal distribution with a peak that traverses the circle with the ensemble average angular velocity but with gradually increasing spread as the phase variable diffuses. We characterize the spread at any instant by the circular standard deviation σ_θ of the phase distribution.

In our simulations using an ensemble of N sample paths we estimate the spread by the sample circular standard deviation $s_\theta = \sqrt{-2 \ln |\frac{1}{N} \sum_{k=1}^{N} e^{i\theta_k}|}$. This definition of the circular standard deviation has the desirable property that in the case of a wrapped normal distribution it gives the same value as the standard deviation of the underlying normal distribution on $(-\infty, \infty)$ that was wrapped around the circle (and so σ_θ ranges from 0 for a Dirac delta distribution to ∞ for a uniform distribution around the circle)(Fisher, 1993). However s_θ is not an unbiased estimator of σ_θ. For a phase variable θ uniformly distributed around the circle, drawing N samples results in an expectation value for the estimator of $\langle s_\theta \rangle \approx 0.152 \ln N + 1.68$ as shown in Figure 5.1. Therefore in our numerical simulations of $N = 1024$ sample paths, a measured value for s_θ of about 2.7 corresponds to a uniform distribution of phase.

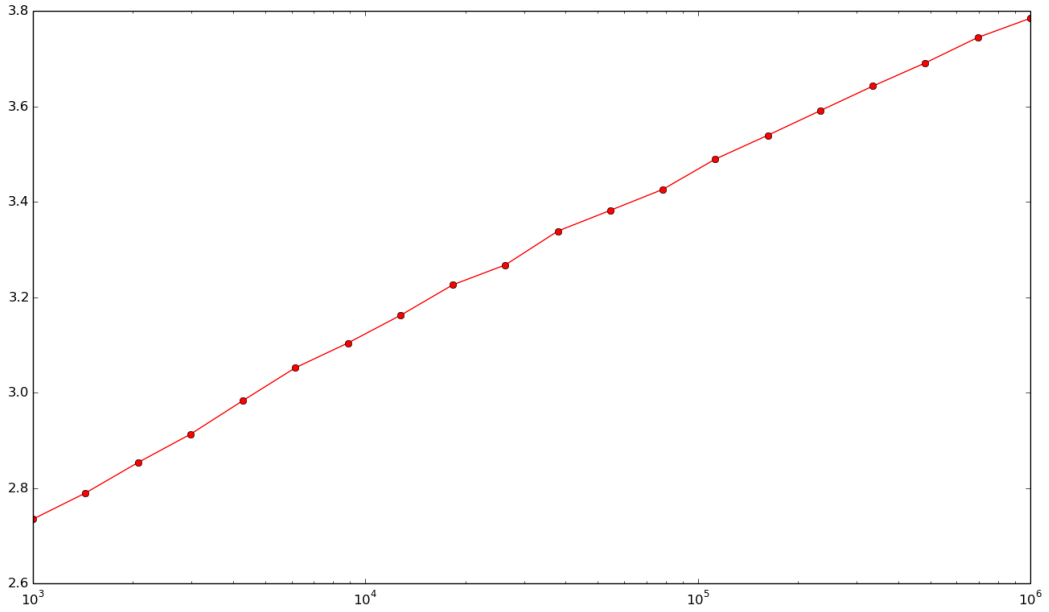


Figure 5.1: $\langle s_\theta \rangle$ for various samples sizes N with all samples drawn from a uniform phase distribution

Stationary probability distribution of the phase

Depending on the system, the limiting stationary distribution of the phase as $t \to \infty$ may not be uniform around the circle. Where the phase space flow is fast at some angles and slow at others, the system is more likely to be found in a slow region at any moment and so the limiting stationary distribution of the phase variable has maxima in these locations.

An interesting consequence of this is that in some systems the circular standard deviation of the phase may not increase monotonically with time, if the starting state has probability concentrated in a relatively fast region.

Probability distribution of amplitude

After normal form transformation, the amplitude variable r is defined up to an arbitrary linear rescaling. As discussed in Section 3.4.1 there are different options for scaling the amplitude of the simplified system and we choose the scaling that makes the coefficient of the cubic term -1 rather than choosing to preserve the numerical value of the amplitude.

To compare the probability distributions of amplitude at different stages in the pipeline we first normalize the amplitude of each system by dividing by the mean amplitude. In cases where we want to examine the evolution of the amplitude distribution through the bifurcation, we want to

keep the scale the same in all figures, so that they are comparable. To this end we use the overall mean amplitude of the system across the bifurcation to normalize the amplitude variable.

In simulation we compare the histograms of the probability distributions, estimated using 1024 sample paths.

5.1.3 Temporal properties of phase and amplitude dynamics

Probability distribution of the period of oscillation

In simulation we report both the mean period and a histogram of the distribution of periods estimated using 1024 sample paths.

In fact we prove in full generality below that the distribution of the period is preserved by the normal form transformation.

Normalized autocorrelation of amplitude

This indicates the "memory" of the amplitude time series, at different lag time intervals τ. As in Chapter 2, to estimate the normalized autocorrelation function we first normalize each time series to a mean of 0 and standard deviation of 1, then compute the cross-correlation of the series with itself applying unbiased normalization,

$$\hat{R}_{xx,\text{unbiased}}(m) = \frac{1}{N-m} \sum_{n=0}^{N-m-1} x_{n+m} x_n \quad (m \geq 0), \tag{5.1}$$

where m is the lag expressed as number of samples (Orfanidis, 1996). Since we are using samples of finite length to estimate the autocorrelation of an (approximately stationary) infinite time series, estimates are necessarily less accurate for longer lags as there are less pairs of points to convolve separated by that lag. Therefore we report autocorrelation for lag times from $\tau = 0$ up to one quarter of the simulated time series length.

Power spectral density of amplitude

As the power spectrum is related to the autocorrelation function by a Fourier transform it contains the same information. But some differences between time series are more easily visible in frequency space, as will be seen below. To estimate the power spectral density (PSD) of the amplitude time series we first subtract any linear trend and normalize the amplitude to a standard

deviation of 1 (that is, normalize to a total power of 1) before applying the Welch algorithm to estimate the power spectrum using 8 Hamming windows with 50% overlap. As the amplitude has been defined by the Hilbert transform of the oscillating signal, which removes higher frequencies, the power spectra estimated here are only meaningful for frequencies lower than the main oscillation frequency.

Distribution of mean reversion times of amplitude

Close to the bifurcation where stability is weak, the amplitude wanders away from its average value for extended periods of time before returning. We characterize this temporal behavior by estimating the mean first return time to the mean amplitude.

The mean amplitude of the process is first estimated by the sample mean over all times and all sample paths in the simulated ensemble. For each simulated sample path we then compute all time intervals between successive returns to the mean. If the actual time of an individual crossing of the mean falls between two integration time steps, we use linear interpolation to estimate the time of the crossing. To ignore spurious results from discrete time approximation we limit attention to excursions further than 0.05 standard deviations from the mean.

5.1.4 Proof that the distribution of periods is preserved

We first prove that the distribution of the stochastic period (that is the distribution over all sample paths of the time intervals of all individual orbits of the phase around the circle) is unchanged by the initial linear transformation and the normal form transformation (steps 1 and 2 in the transformation pipeline).

The initial linear transformation is invertible (it is defined as the similarity transformation to Jordan real form) Because the normal form transformation (step 2) is near-identity (it is equal to the identity to first order) so it is invertible within a neighborhood of the origin in phase space.

Since these two transformations are invertible they are one-to-one. Hence for each realization of the noise process, the system passes its starting point at the same moment in the original and the transformed system. Thus for every sample path the time interval of each individual orbit of the circle

is preserved by the transformation. Therefore the probability distribution of the period (over all orbits of all sample paths) is preserved by the transformation.

As an important corollary the mean frequency of the oscillator is preserved by these transformations.

5.1.5 Framework for parallel stochastic simulation

To characterize a stochastic dynamical system statistically through simulation, even at a single point in parameter space, is computationally intensive. This is because a single simulation run can give (an approximation of) a single sample path, but an ensemble of one thousand or more sample paths may be needed to estimate the statistics described above that characterize the oscillations.

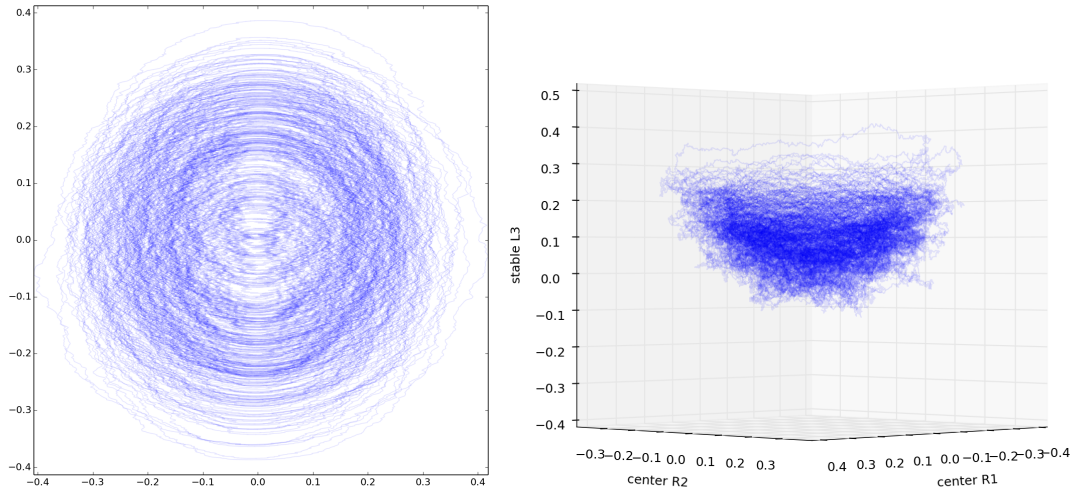
To facilitate this, a new software platform was developed to make it easy to conduct massively parallel simulations of stochastic dynamical systems and then to apply analyses to the resulting time series in parallel. This simulation and analysis platform has been released as a free software project and is available at https://pypi.python.org/pypi/nsim

It supports high performance computing clusters, cloud services or a single workstation with many processing cores, parallelizing both the initial simulation of sample paths and the subsequent computation of time series statistics.

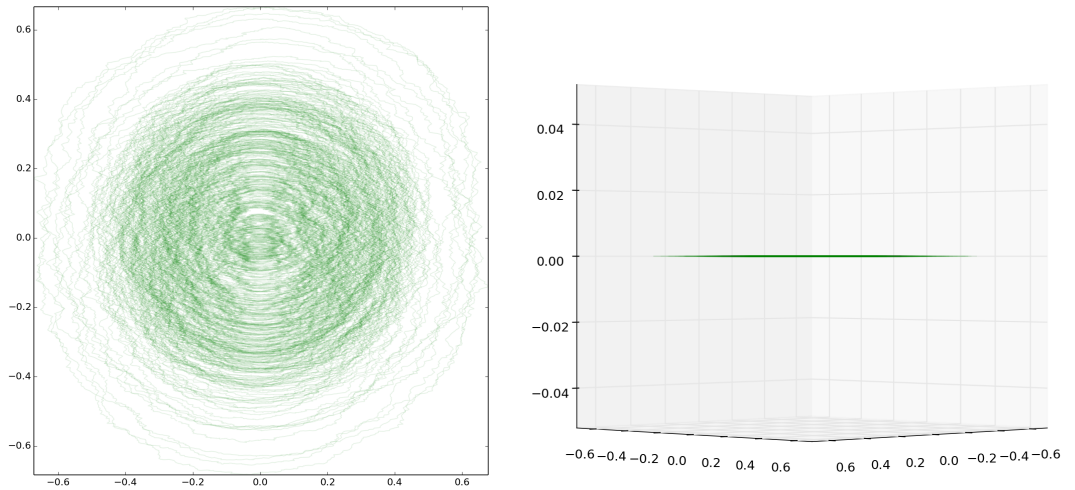
Integration of the stochastic differential equations was performed using the strong order 1.0 Stratonovich Runge-Kutta algorithm SRS1 of Rößler (2010). Polar equations were put in Cartesian form for numerical integration. The implementation is detailed in Appendix E.

5.2 Symmetrical example system

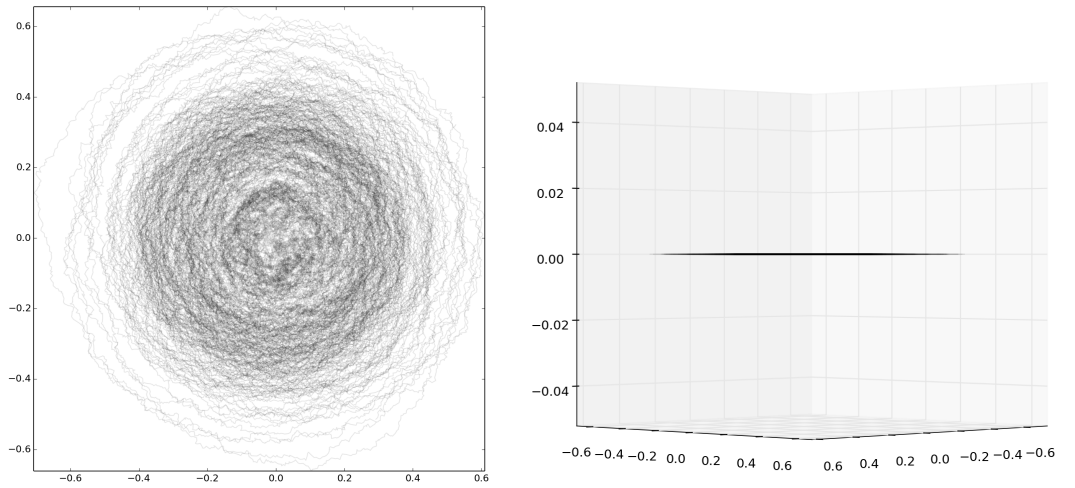
Before introducing more complicated systems we illustrate these statistics using the simple symmetrical 3D system which we have used as the first example through Chapters 3 and 4.



A Original 3D system



B 2D reduced system



C Final averaged system

Figure 5.2: Simple example system: Orbits in phase space at the deterministic bifurcation point $\epsilon = 0$, top view (left) and side view (right), for the 3D original, 2D reduced and final averaged system respectively, each showing 180 seconds of a single sample path.

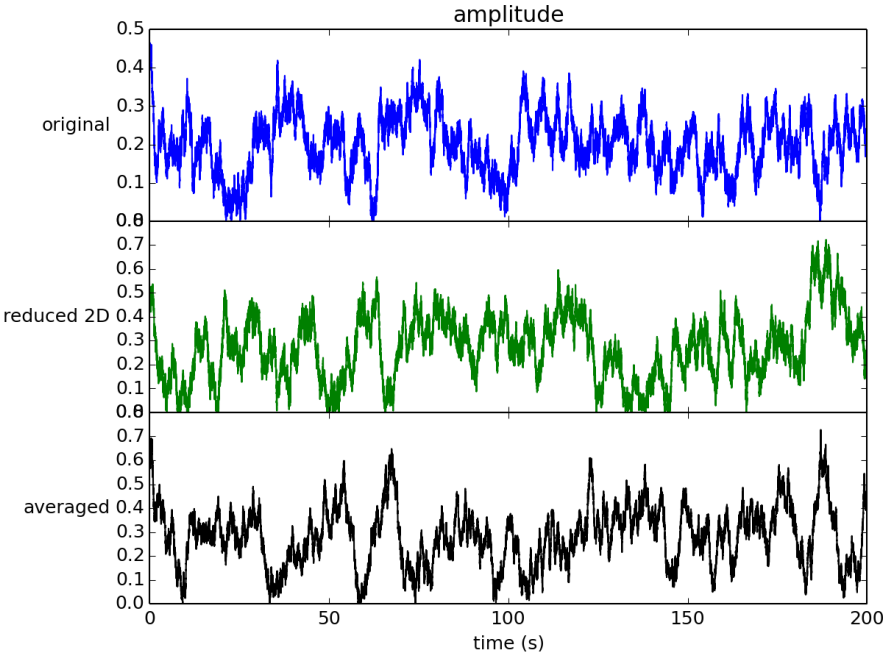


Figure 5.3: Results for simple example system: Example amplitude time series at $\epsilon = 0$, for the original three-dimensional system (blue), after normal form transformation and reduction to two dimensions (green) and final averaged system with rescaled radius (black).

5.2.1 Amplitude and phase distributions

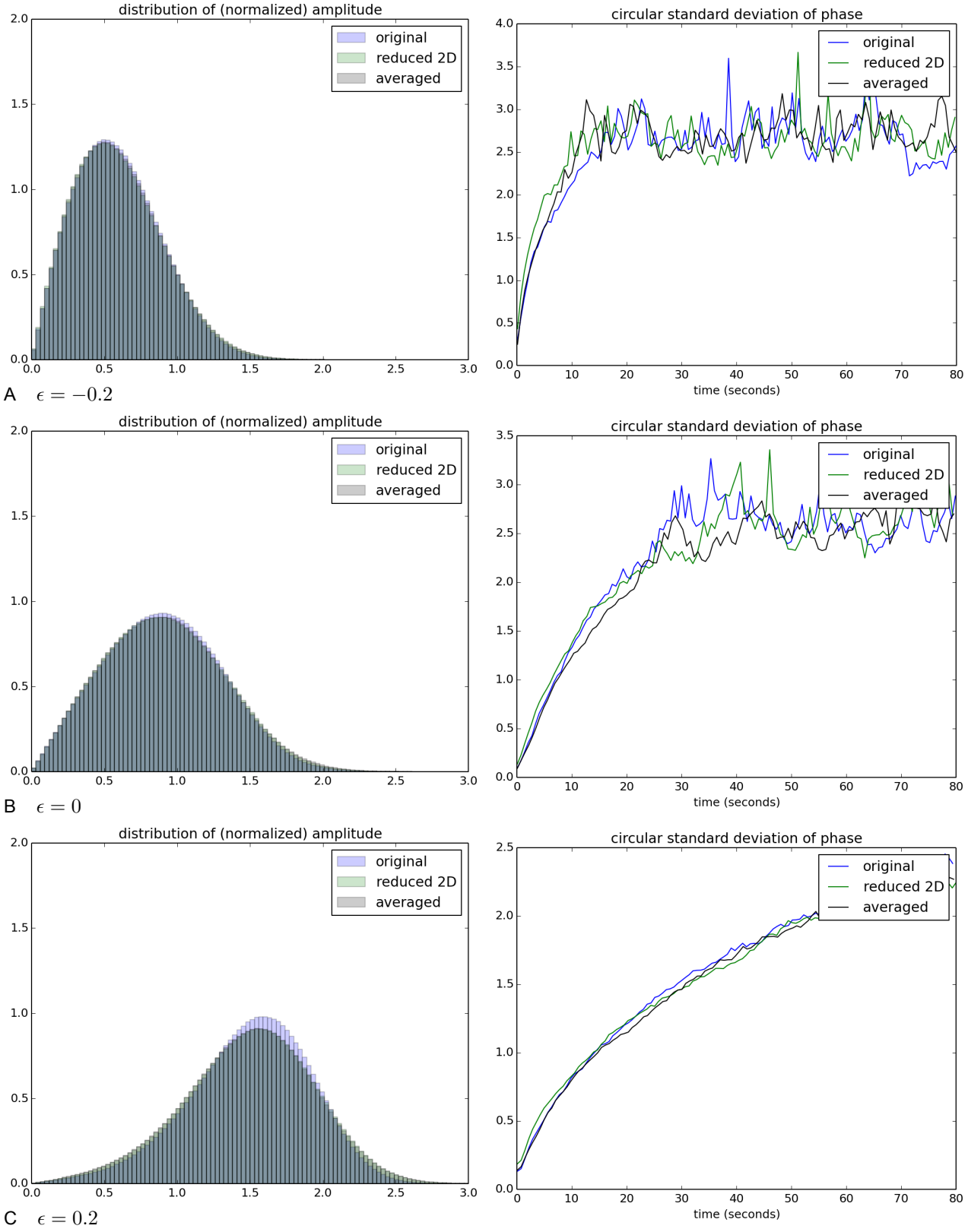


Figure 5.4: Results for simple example system: Normalized amplitude distribution (left) and phase diffusion (right) before, at and after the deterministic bifurcation point $\epsilon = 0$, superimposing the original 3D system (blue), 2D transformed system (green) and final averaged system (black).

5.2.2 Temporal properties

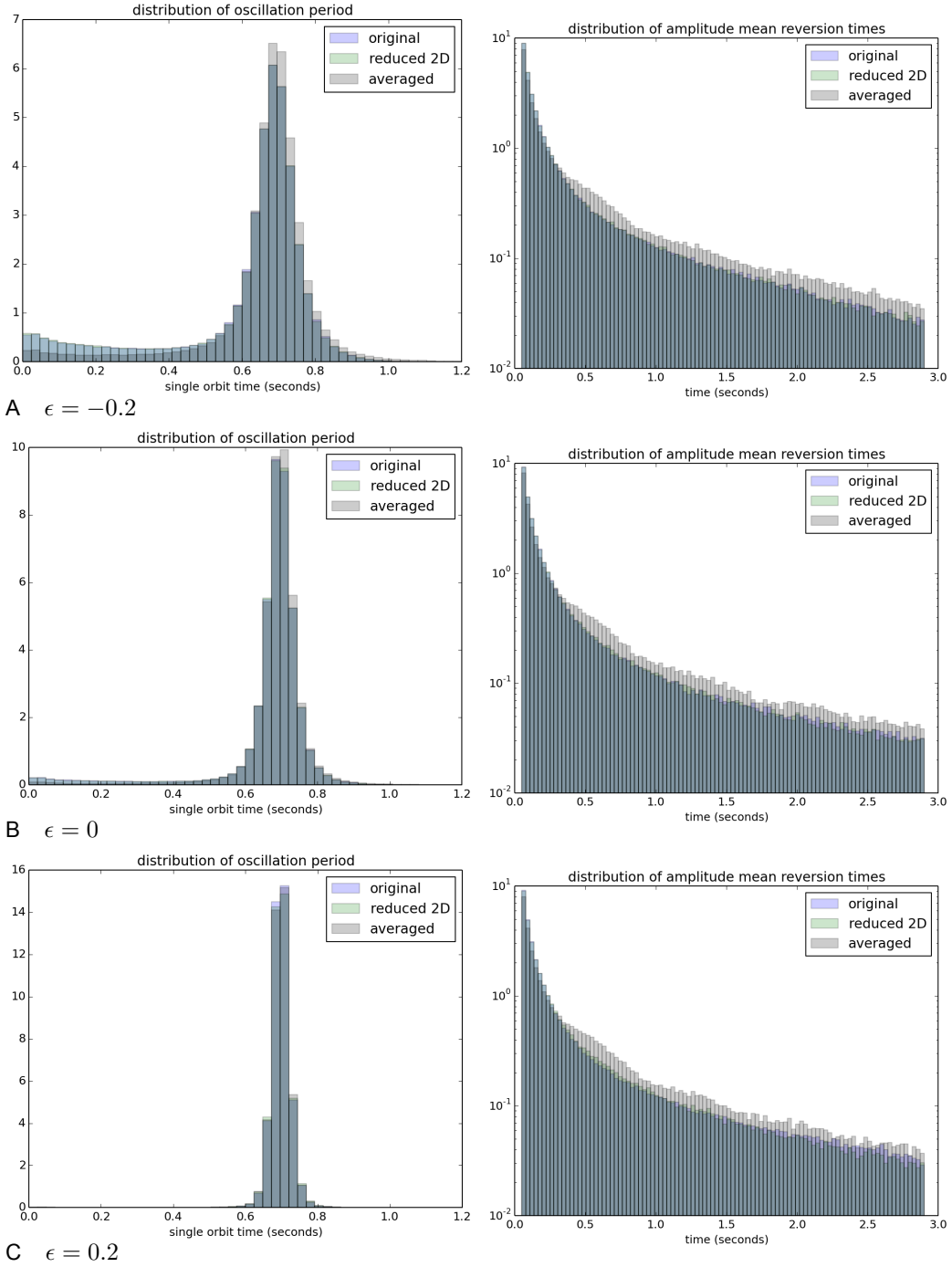


Figure 5.5: Results for simple example system: Distribution of stochastic period (left) and mean reversion time (right) before, at and after the deterministic bifurcation point $\epsilon = 0$, superimposing the original 3D system (blue), 2D transformed system (green) and final averaged system (black).

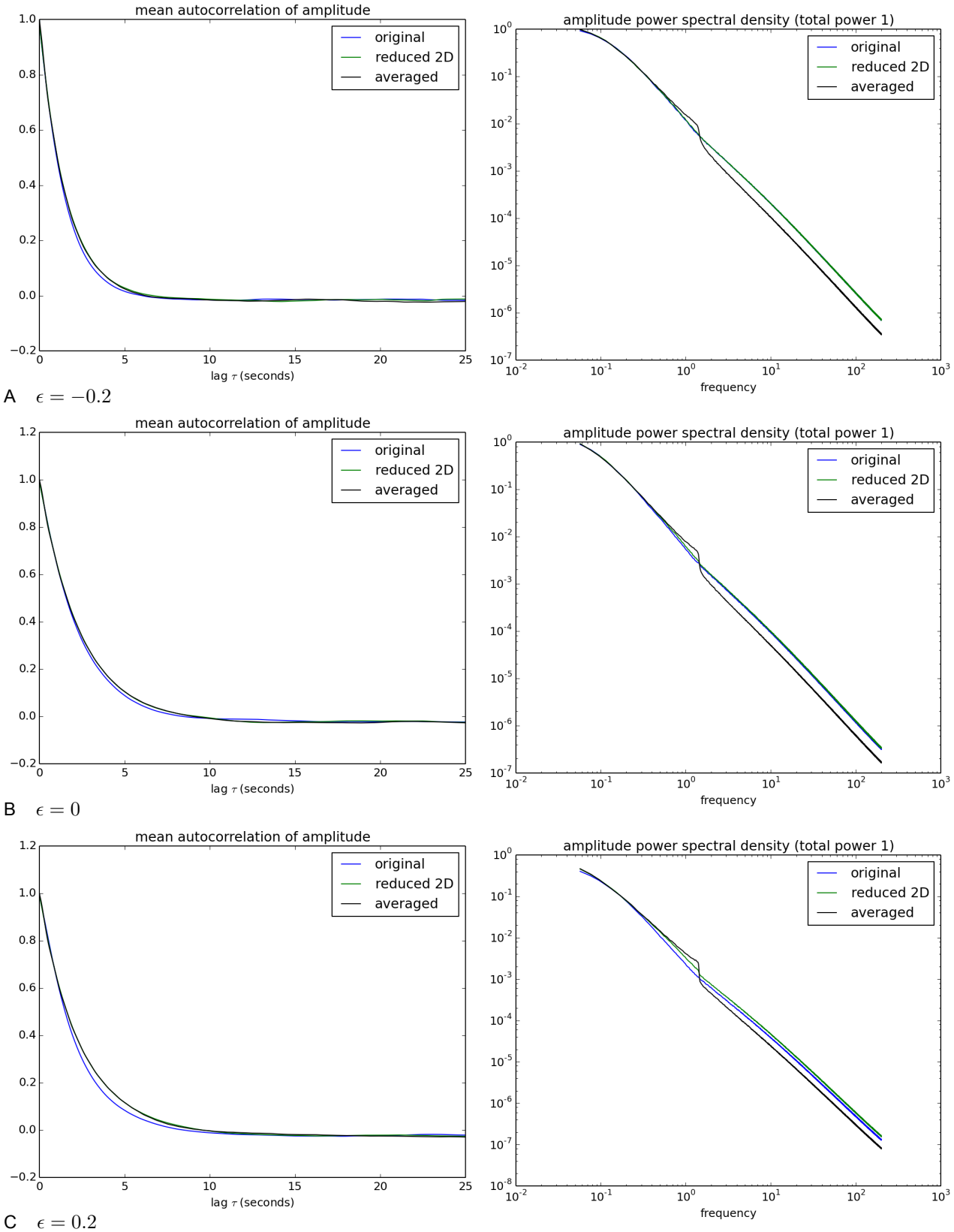


Figure 5.6: Results for simple example system: autocorrelation of amplitude (left) and power spectral density (right) before, at and after the deterministic bifurcation point $\epsilon = 0$, superimposing the original 3D system (blue), 2D transformed system (green) and final averaged system (black).

5.3 Asymmetrical, non-orthogonal test system

The example system in Section 5.2 is minimal. In order to be minimal it has symmetries that make it unrepresentative of the general case, specifically:

- Deterministic terms already have circular symmetry (equivariance) before transformation;
- The initial linear transformation is trivial, so we can not examine changes to oscillation properties due to a non-orthogonal linear transformation.

This section introduces a non-trivial three-dimensional test system deliberately constructed so that emerging oscillations are asymmetrical and eigenvectors of the original system's Jacobian matrix at the bifurcation are far from orthogonal. This is the original system labeled **O**:

$$\begin{aligned}
dx_1 &= (311.479x_1 - 2201.82x_2 - 564.635x_3 - 20.0855x_1^2 \\
&\quad + 465.343x_2^2 + 65.9025x_1x_3 + 52.214x_1x_2 - 76.4356x_2x_3 - 54.2295x_3^2)\,dt \\
dx_2 &= (-2.21885x_1 + 30.7184x_2 + 6.63859x_3 + 0.255836x_1^2 \\
&\quad - 6.98162x_2^2 - 0.933232x_1x_3 - 1.40187x_1x_2 + 2.24106x_2x_3 + 0.87356x_3^2 - \alpha)\,dt \\
dx_3 &= (195.124x_1 - 1312.67x_2 - 347.197x_3 - 12.4969x_1^2 \\
&\quad + 255.026x_2^2 + 40.7643x_1x_3 - 53.1318x_2x_3 + 33.9785x_1x_2 - 33.4977x_3^2)\,dt + \sigma\,dW.
\end{aligned} \tag{5.2}$$

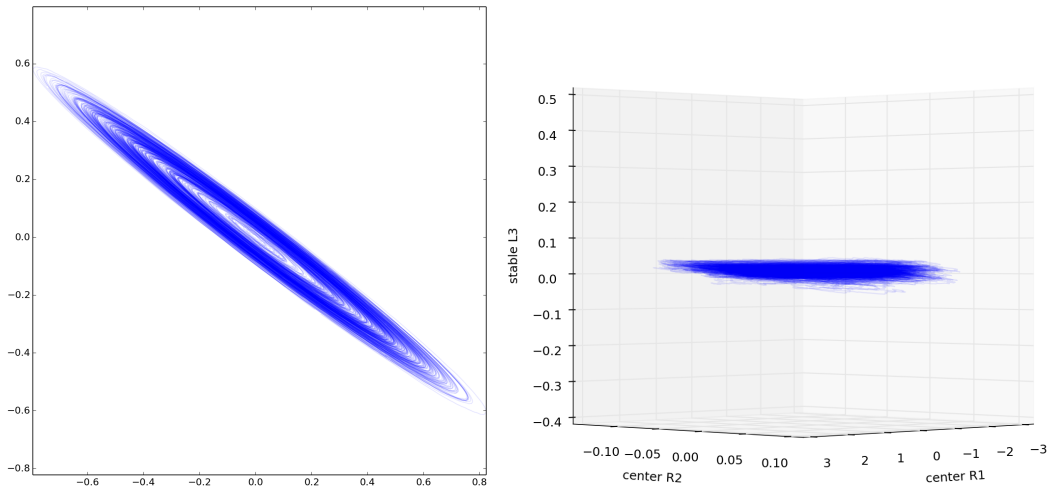
For this system the ratio of largest to smallest singular values of the Jacobian at bifurcation is 3581.21 so it is far from orthogonal. The details of its construction and derivation of the transformations below are presented in Appendix B. Applying our method to this test system, after **steps 1, 2 and 3** (initial linear transformation, normal form transformation and reduction to 2D), we have the intermediate

system labeled **R**:

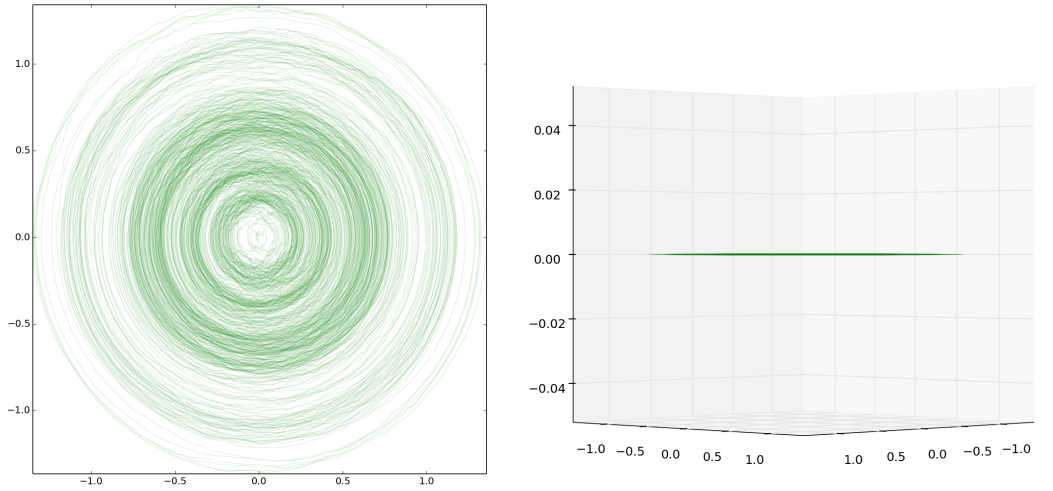
$$
\begin{aligned}
dr = (&-0.135078r^3 - \alpha(9.86728\sin\theta - 0.788631\cos\theta) \\
&+ \alpha r(0.739492 + 0.0309986\sin 2\theta - 0.0344784\cos 2\theta))\,dt \\
&+ (2.48988\sin\theta - 0.356561\cos\theta + r(0.321319 - 0.0104069\sin 2\theta - 0.00338031\cos 2\theta) \\
&+ r^2(0.135457\sin\theta - 0.0164455\cos\theta + 0.0000318211\sin 3\theta + 0.000119974\cos 3\theta))\sigma \circ dW \\
d\theta = (&69 - 0.454487r^2 + \frac{\alpha}{r}(0.788631\sin\theta - 9.86728\cos\theta) \\
&+ \alpha(2.38805 + 0.0309986\cos 2\theta + 0.0344784\sin 2\theta))\,dt \\
&+ (\frac{1}{r}(0.356561\sin\theta + 2.48988\cos\theta) + 1.0797 - 0.0104069\cos 2\theta + 0.00338031\sin 2\theta \\
&+ r(0.456461\sin\theta - 0.0577074\cos\theta - 0.000119974\sin 3\theta + 0.0000318211\cos 3\theta))\sigma \circ dW.
\end{aligned}
\tag{5.3}
$$

Finally after **step 4** (averaging around the cycle and rescaling of radius), we have the transformed system in standard form, labeled **A**:

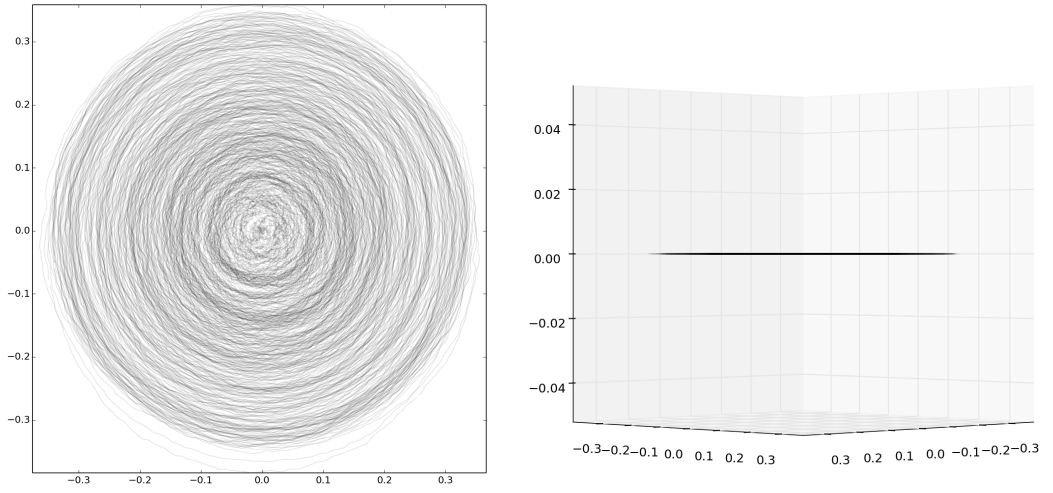
$$
\begin{aligned}
dR = (&-R^3 + r(0.739492\alpha + 0.0906116\sigma^2) + \frac{1}{R}0.213647\sigma^2)\,dt \\
&+ (0.653677 + 0.341485R^2)\sigma \circ dW_A \\
d\theta = (&69 + 2.38805\alpha - 0.17714\sigma^2 - 3.36464R^2)\,dt \\
&+ 1.42144R\sigma \circ dW_A + (\frac{1}{R}0.653677 + 0.906326R)\sigma \circ dW_B.
\end{aligned}
\tag{5.4}
$$



A Original 3D system



B 2D reduced system



C Final averaged system

Figure 5.7: Asymmetric non-orthogonal test system: Orbits in phase space at the deterministic bifurcation point $\epsilon = 0$, top view (left) and side view (right), for the 3D original, 2D reduced and final averaged system respectively, each showing 40 seconds of a single sample path. Note that an orbit in the original system is a sharp ellipse: the top figure is not an oblique view of a circle.

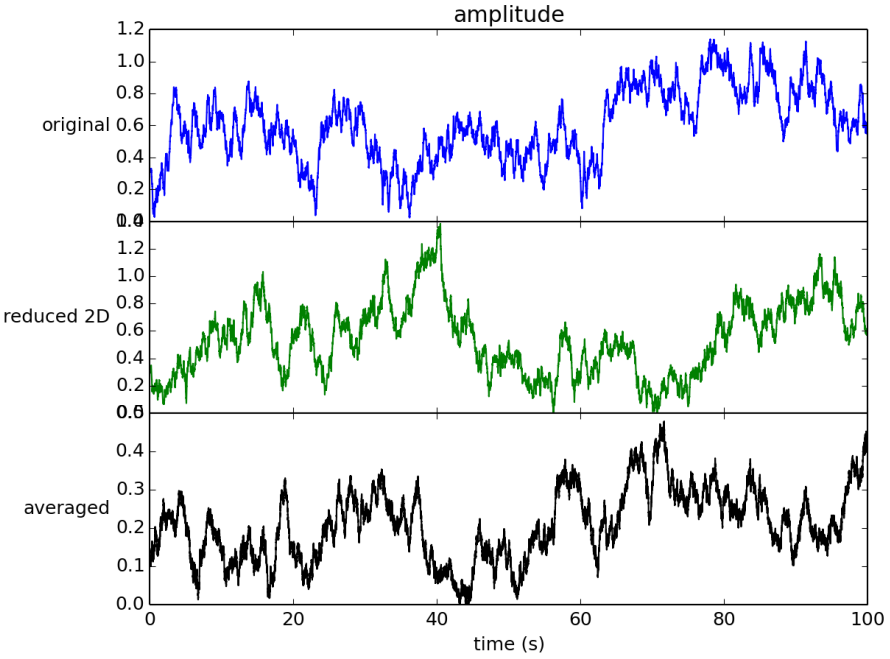


Figure 5.8: Results for asymmetric non-orthogonal system: Example amplitude time series at $\epsilon = 0$, for the original three-dimensional system (blue), after normal form transformation and reduction to two dimensions (green) and final averaged system with rescaled radius (black).

5.3.1 Test system: amplitude and phase distributions

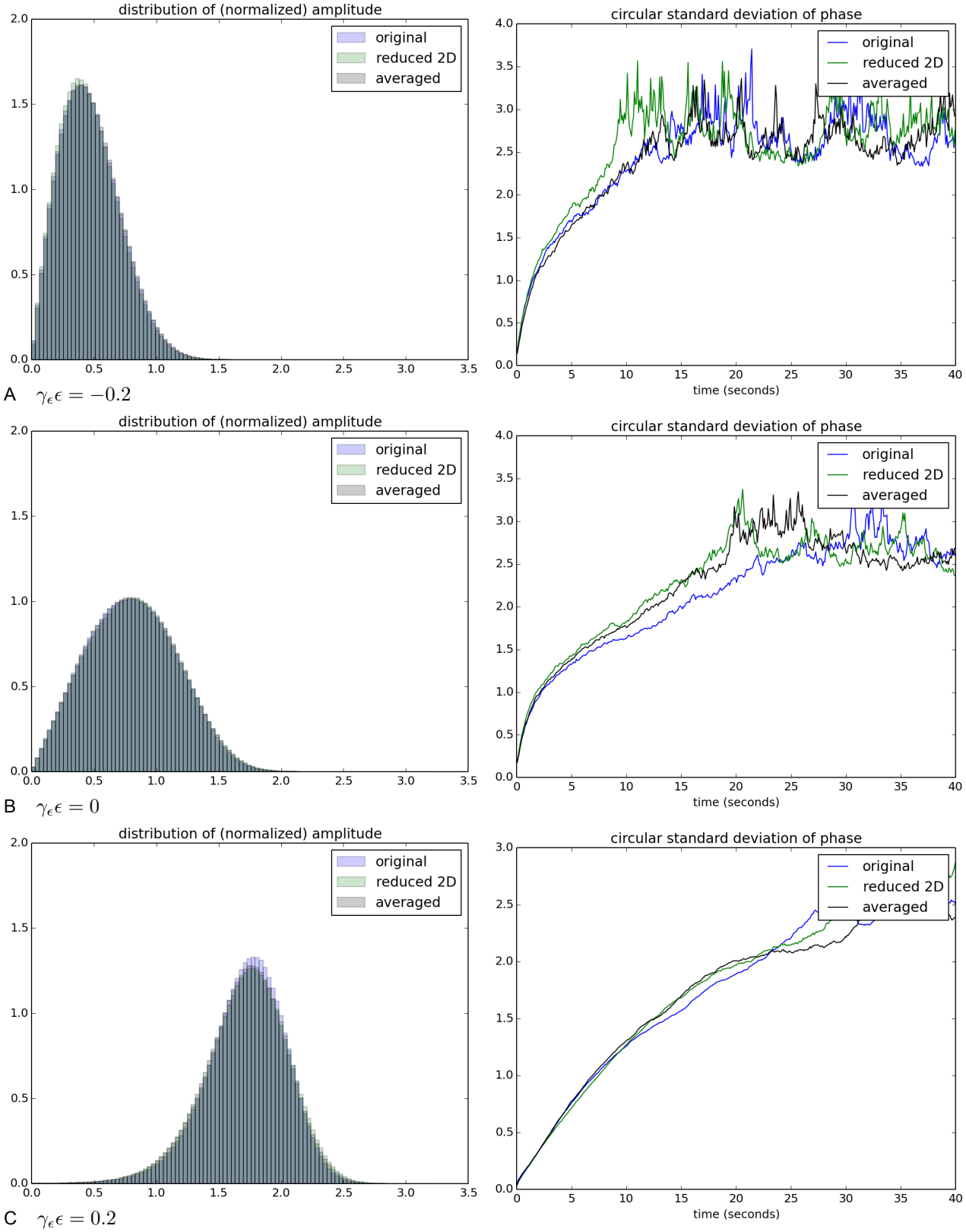


Figure 5.9: Results for asymmetric non-orthogonal system: Normalized amplitude distribution (left) and phase diffusion (right) before, at and after the deterministic bifurcation point $\epsilon = 0$, superimposing the original 3D system (blue), 2D transformed system (green) and final averaged system (black).

5.3.2 Test system: temporal properties

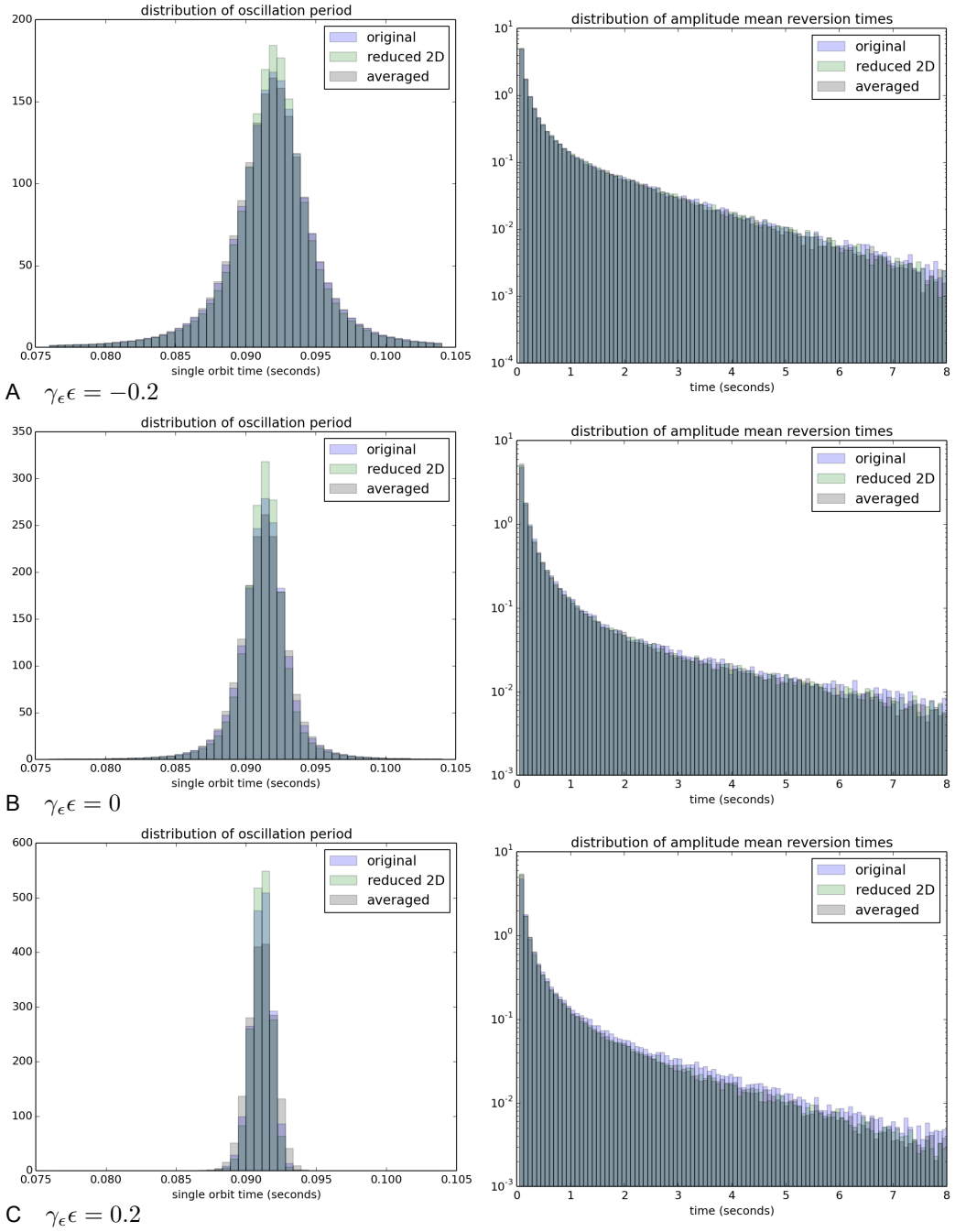


Figure 5.10: Results for asymmetrical, non-orthogonal test system: Distribution of stochastic period (left) and mean reversion time (right) before, at and after the deterministic bifurcation point $\epsilon = 0$, superimposing the original 3D system (blue), 2D transformed system (green) and final averaged system (black).

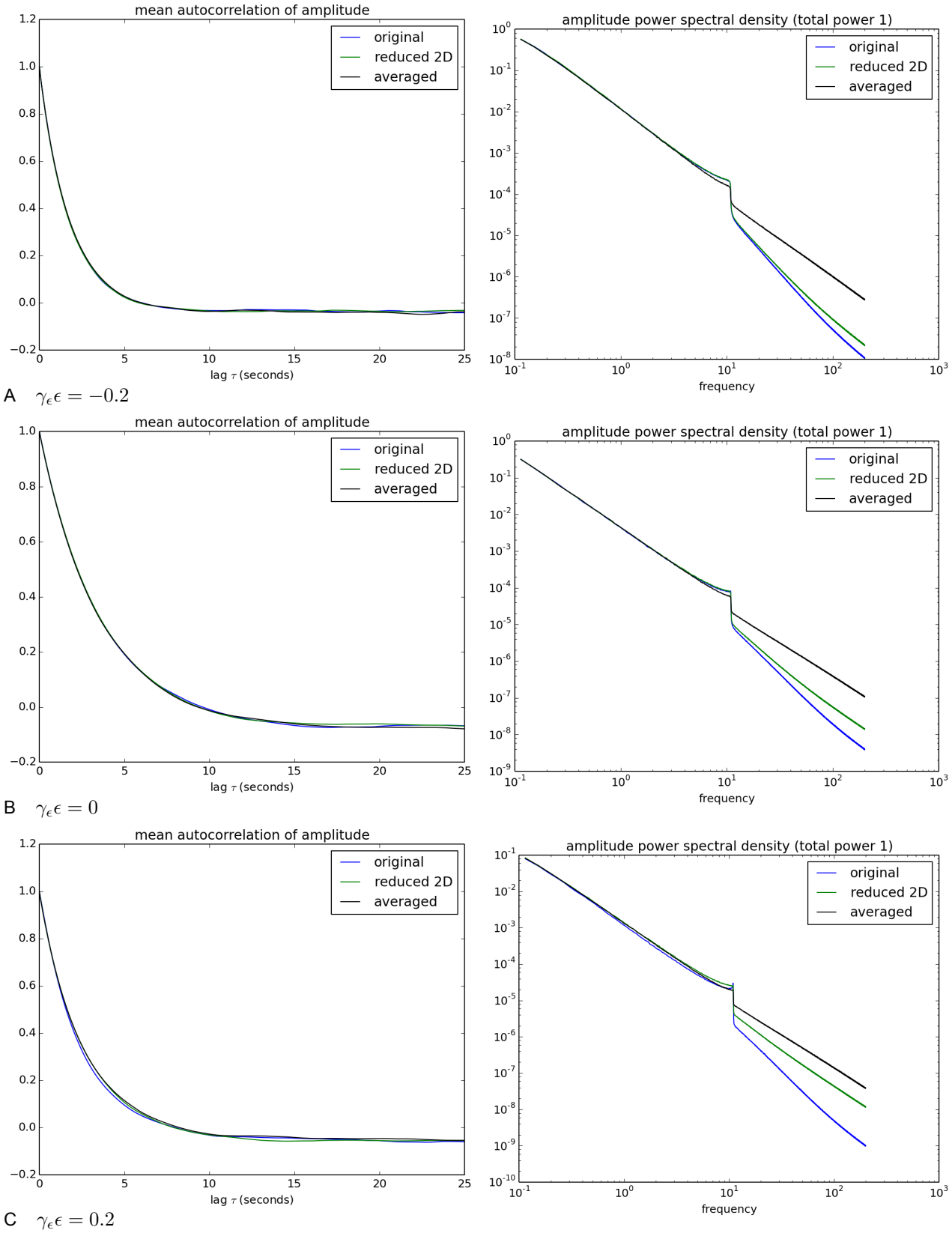


Figure 5.11: Results for asymmetrical, non-orthogonal test system: autocorrelation of amplitude (left) and power spectral density (right) before, at and after the deterministic bifurcation point $\epsilon = 0$, superimposing original 3D system (blue), 2D transformed system (green) and averaged system (black).

As we first applied the Hilbert transform (implemented internally by discrete Fourier transform and its inverse) to extract the amplitude envelope, attenuating frequencies higher than the oscillation, we do not interpret the power spectral results for frequencies greater than the dominant oscillation frequency, in this case 11 Hz.

5.4 Summary of results

All of the statistical properties examined were approximately preserved by the initial linear transformation and near-identity normal form transformation.

Excepting the distribution of mean reversion times, all properties were approximately preserved through the entire pipeline. In the case of the distribution of mean reversion times, averaging preserves the distribution at intermediate time scales, but not for short time scales less than the period. The prevalence of short excursions from the mean is significantly reduced by averaging. This discrepancy at short time scales is reasonable, as the point of averaging is to derive an approximation valid only over longer time scales, where the phenomena of interest are slower than individual oscillations.

Chapter 6

Application to Jansen-Rit model of a cortical region

Recalling the results of Chapter 2, when simulating the Jansen-Rit neural mass model we observed effects of noise close to its branch of supercritical Hopf bifurcations that were unexpected:

- Long autocorrelation times were evident near a Hopf bifurcation point when noise perturbation was to the spiny stellate cell population (H3u), but not when noise perturbed the pyramidal cell population (H3p), even though the spiny stellate direction was better aligned in state space with the center eigenspace of the bifurcation.
- Increased variance and increased autocorrelation of the time series were both used as possible statistical indicators of the loss of linear stability, but scenario H3u showed lower variance at the same time as greatly increased autocorrelation compared to H3p.

The same contrast was seen between Hopf points H1 and H2. In the discussion of Chapter 2 we hypothesized that the cause of these features might be the curvature of the center manifold, noting that multiplicative noise results from this curvature, but that we did not know its magnitude or importance.

We also observed that oscillations shifted between high and low amplitude on a time scale of 5-10 seconds, with the amplitude envelope having a balloon like form, and hypothesized that this could also be due to significant multiplicative noise, as is seen when comparing time series of the scalar

additive and multiplicative Ito processes,

$$dx = (\epsilon - ax^2)x\,dt + \sigma\,dW, \qquad \text{and} \tag{6.1}$$

$$dx = (\epsilon - ax^2)x\,dt + \sigma x\,dW. \tag{6.2}$$

But it was by applying the method of Chapters 4 and 5 considering the simplified system in transformed coordinates that the reasons for these behaviors became clear. In fact our hypotheses were wrong. The magnitude of multiplicative noise, and hence curvature of the center manifold, were insignificant in this system. Instead it is the linear term that is key, specifically the extreme non-normality of the Jacobian in the Jansen-Rit system.

This is explained in detail below. The pipeline of transformations described in Chapter 4 is used to derive a symmetrical 2D weak approximation to the oscillations near the bifurcation. From this standard form we can then see directly and quantitatively the relative importance of additive and multiplicative noise effects, and also view the differences in noise-induced effects between scenarios H3p and H3u.

In Chapter 2 we also noted that it was the reference plane normal to the stable eigenspace that mattered when assessing the effect of noise direction on the oscillations, not the alignment of noise with the center eigenspace. By considering the transformation of the system to its normal form coordinates it is straightforward to see why this is, as discussed below.

Finally we discuss the implications for modeling of biological systems and the interpretation of autocorrelation in neuroimaging time series.

6.1 The Jansen-Rit equations as SDE

The Jansen-Rit equations, extended to allow separate input to pyramidal and spiny stellate populations, were introduced in Eqs (2.1)–(2.7). Expressed as a first order Ito SDE in the variables $\mathbf{v} = (v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8)^T$ this is,

$$\mathbf{dv} = \mathbf{f}(\mathbf{v})\,dt + G\,\mathbf{dW}(t), \tag{6.3}$$

where

$$\mathbf{f}(\mathbf{v}) = \begin{bmatrix} v_5 \\ v_6 \\ v_7 \\ v_8 \\ H_e\kappa_e(\gamma_1 S(v_2 - v_3) + \langle u \rangle) - \kappa_e^2 v_1 - 2\kappa_e v_5 \\ H_e\kappa_e(\gamma_2 S(v_1) + \langle p \rangle) - \kappa_e^2 v_2 - 2\kappa_e v_6 \\ H_i\kappa_e\gamma_4 S(v_4) - \kappa_i^2 v_3 - 2\kappa_i v_7 \\ H_e\kappa_e\gamma_3 S(v_2 - v_3) - \kappa_e^2 v_4 - 2\kappa_e v_8 \end{bmatrix}, \quad G\,\mathbf{dW}(t) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ H_e\kappa_e\sigma_u & 0 \\ 0 & H_e\kappa_e\sigma_p \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} dW_1(t) \\ dW_2(t) \end{bmatrix},$$

(6.4)

$$S(v) = \frac{2e_0}{1 + \exp[\rho_1(\rho_2 - v)]}.$$

for the case of uncorrelated fluctuations in pyramidal and spiny stellate input. If modeling fully correlated fluctuations in input to the two populations, the noise term is instead

$$G\,\mathbf{dW}(t) = H_e\kappa_e(0, 0, 0, 0, \sigma_u, \sigma_p, 0, 0)^{\mathrm{T}}\, dW_1(t). \tag{6.5}$$

Here using the standard parameter values of Jansen and Rit (1995), as listed in Table 2.1, we first set $\langle u \rangle = 0$ and $\sigma_u = 0$ to focus on the supercritical Hopf bifurcation point labeled H1 in Section 2.3.1. This is the same point in parameter space observed by Jansen and Rit (1995) and analyzed by Grimbert and Faugeras (2006). As shown in Section 2.3.1, when $\langle u \rangle = 0$, increasing $\langle p \rangle$ takes the system through a supercritical Hopf bifurcation at $\langle p \rangle = p_c \approx 89.8\,\mathrm{s}^{-1}$. Near the bifurcation point, define the small parameter ϵ by $\langle p \rangle = p_c + \epsilon$ so that the bifurcation occurs at $\epsilon = 0$. $\mathbf{f}$ can then be considered a function of $\mathbf{v}$ and ϵ.

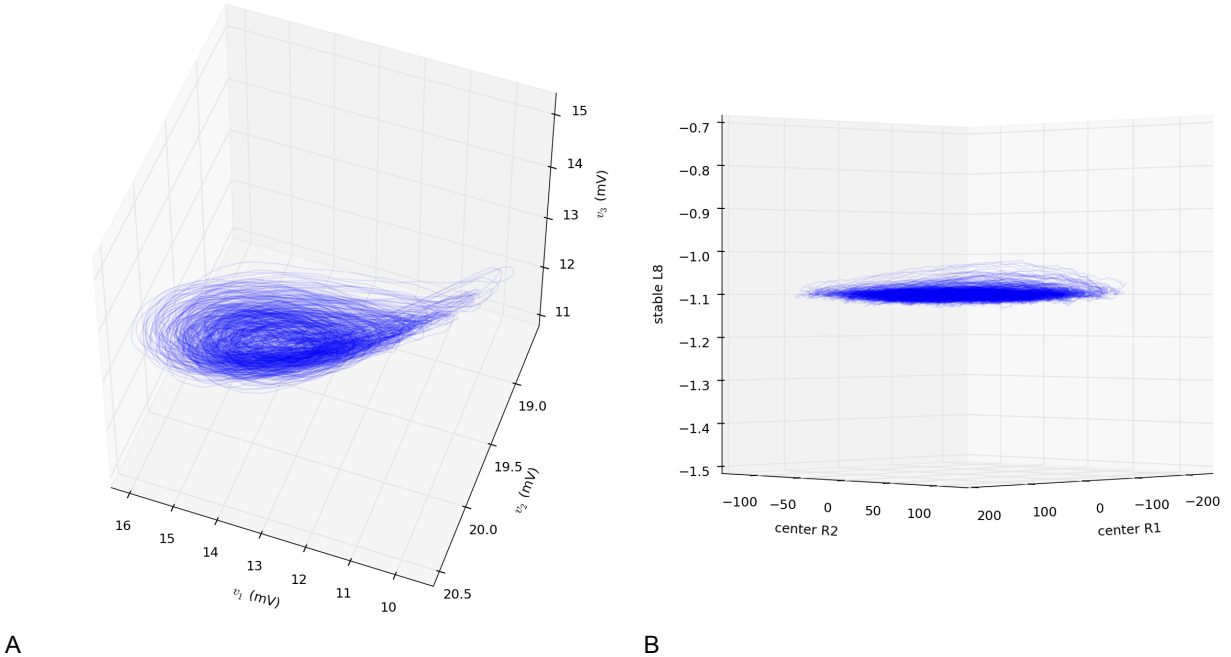
6.1.1 Computing the bifurcation point

In Chapter 2 both the value p_c in parameter space where the bifurcation occurs and the location in state space of the equilibrium point were found approximately for each of points H1, H2 and H3 using numerical continuation in the package MATCONT. But in the Jansen-Rit system it is possible to solve directly for the bifurcating equilibrium point $\mathbf{v}_c = (v_{c1}, v_{c2}, v_{c3}, v_{c4}, v_{c5}, v_{c6}, v_{c7}, v_{c8},)^T$ and bifurcation

point p_c for any given value of $\langle u \rangle$. This derivation is given in Appendix A.1. Since p_c is a continuous, monotonic function of $\langle u \rangle$ near H3, the location of H3 (defined as the point where the Hopf branch intersects the line $\langle u \rangle = \langle p \rangle$) can also be computed to arbitrary precision. For the case of supercritical Hopf point H1, we obtain

$$p_c = 89.8291\,\mathrm{s}^{-1}\,, \qquad \mathbf{v}_c = (13.2082, 20.1649, 13.4254, 3.30206, 0, 0, 0, 0)^{\mathrm{T}}\ \mathrm{mV}. \tag{6.6}$$

Figure 6.1 shows two different projections to 3D of the same 8D orbit in the original Jansen-Rit system at this point H1. This illustrates how the 2D center manifold near which the oscillations emerge embedded in the 8D space is flat in some dimensions and curved in others, resulting in oscillations that are non-elliptical.



A B

Figure 6.1: Two choices of 3D projection of a Jansen-Rit orbit at supercritical Hopf point H1. Showing 40 seconds of the same sample path in 8 dimensions, (A) projected to the first 3 dimensions (v_1, v_2, v_3), (B) projected to the center right eigenspace and the left eigenspace of the weakest stable eigenvalue.

6.2 Applying the method to the Jansen-Rit system

We apply the method of Chapter 4 to examine the system close to bifurcation in a normal form coordinate system, a choice of coordinates which maps the oscillations (in the absence of noise) to circles in a flat 2D plane. These steps are automated by the accompanying Mathematica package which can be applied to any system near a supercritical Hopf bifurcation (Appendix D). The detailed equations at each step when transforming the Jansen-Rit system at H1 are presented in Appendix C. Points H2 and H3 are treated similarly. Here we note some specific challenges in applying the method to the Jansen-Rit system and focus on comparing the results in standard form for scenarios H3p, H3u, H1 and H2 of Chapter 2.

The Jansen-Rit system is a good candidate for this method because the stable eigenspace is very stable, with the weakest stable eigenvalue having real part $-\lambda \approx -80.14$ so that the criterion $\frac{\sigma^2}{\lambda^{1.5}} \ll 1$ (discussed in Section 4.4.3) will be met for the noise intensity $\sigma_p \approx 0.54$ that was used in the simulations of Chapter 2. This gives some assurance that the two-dimensional approximation will be useful.

During **step 1** (translation to the origin and linear transformation) note that the Jansen-Rit system is highly ill-conditioned, including for eigenvector computation, as its Jacobian matrix $A = D\mathbf{f}(\mathbf{v})|_{\mathbf{v}_c}$ has a ratio of largest to smallest singular values of more than 40000 in the physiological region of parameter space. To improve precision we apply the method of Parlett and Reinsch (1969) to incorporate in the linear transformations an initial rescaling by the diagonal matrix $T = \text{diag}(\frac{1}{16}, \frac{1}{32}, \frac{1}{32}, \frac{1}{32}, 8, 4, 4, 4)$. This reduces the condition number of the Jacobian from 44822 to 17.96. This facilitates computation, though a rescaling does not change the inherent sensitivity of the Jansen-Rit model to its parameters. The transformation matrix G to Jordan real form is then computed by Jordan decomposition of $T^{-1}AT$, so that the overall linear transformation M to be used in step 1 is $M = TG$.

After the translation and linear transformation $\mathbf{v} - \mathbf{v}_c = M\mathbf{x}$, the linearly transformed system,

$$\mathbf{dx} = M^{-1}\mathbf{f}(M\mathbf{x} + \mathbf{v}_c, \epsilon)dt + M^{-1}G(M\mathbf{x} + \mathbf{v}_c) \circ \mathbf{dW}, \tag{6.7}$$

now has a Jacobian B which is in Jordan real form at the origin, with center eigenspace plane given by the first two variables x_1 and x_2. Appendix C.1 shows the full form of Eq. (6.7) calculated for the Jansen-Rit system at H1. At this stage the center manifold of the bifurcation is still curved.

For **step 2** (near-identity normal form transformation) the Jacobian B of the Jansen-Rit model is not diagonalizable over $\mathbb{C}$, having a non-trivial Jordan block with eigenvalue -100. Occurring in such a highly stable eigenspace this will have negligible effect on the oscillations, but it does mean that the semisimple normal form algorithm of previous examples cannot be applied. Instead the inner product normal form style is used (Murdock, 2003), as explained in Section 3.3.2, where $\ker \mathfrak{L}_{B^\dagger}$ rather than $\ker \mathfrak{L}_B$ is used as a complement of $\operatorname{im} \mathfrak{L}_B$. The accompanying Mathematica code automatically makes this selection. Setting the noise intensity σ_p and bifurcation parameter ϵ to zero, the near-identity transformation $\mathbf{x} = \mathbf{y} + \mathbf{q}(\mathbf{y})$ is computed. This transformation is then applied to the full noise-perturbed system Eq. (6.7) using the method of Eq. (4.18) to obtain an 8-dimensional normal form system with noise terms also expressed in the new coordinates.

After **steps 3 and 4** (reduction to 2D, averaging around the cycle and linear rescaling of radius) we obtain a 2D weak approximation in the standard form described in Section 4.5,

$$\begin{aligned} dR &= (-R^3 + (\gamma_\epsilon \epsilon + \gamma_\sigma \sigma^2)R + h\sigma^2 \frac{1}{R})\, dt \quad + \quad (p + qR^2)\sigma \circ dW_A, \\ d\theta &= (\omega + \omega_\epsilon \epsilon + \omega_\sigma \sigma^2 + bR^2)\, dt \quad + \quad s_a R\sigma \circ dW_A \quad + \quad (\frac{p}{R} + s_b R)\sigma \circ dW_B. \end{aligned} \tag{6.8}$$

The above steps produce local approximations that are useful in some neighborhood of the bifurcation point in state space and parameter space, but we did not quantify the size of the neighborhoods where the successive approximations are useful. As a check of validity of the approximations we use simulation of sample paths (similarly to Chapter 5) to verify that the simplified system preserves the statistics of interest, despite the extreme numerical ill-conditioning of the Jansen-Rit system. This is shown in Appendix C.3.

6.3 Results: oscillation parameters

Comparing the constructed scenarios H3p and H3u of Chapter 2, which simulated the same point H3 in parameter space and were identical except that scenario H3p had fluctuation of pyramidal input only, whereas H3u had fluctuation of spiny stellate input only. For scenario H3p the resulting simplified

system is:

$$dR = (-R^3 + (0.0167\epsilon - 0.00290\sigma^2)R + 0.131\sigma^2\frac{1}{R})\,dt \;+\; (0.511 - 0.00142R^2)\sigma \circ dW_A,$$
$$d\theta = (-69.4 - 0.0323\epsilon - 0.00574\sigma^2 + 1.79R^2)\,dt + 0.00267R\sigma \circ \mathrm{d}W_A + (0.511\frac{1}{R} - 0.00919R)\sigma \circ dW_B. \quad (6.9)$$

By comparison in scenario H3u,

$$dR = (-R^3 + (0.0167\epsilon - 0.000210\sigma^2)R + 0.0000990\sigma^2\frac{1}{R})\,dt \;+\; (0.0141 + 0.00833R^2)\sigma \circ dW_A,$$
$$d\theta = (-69.4 - 0.0323\epsilon + 0.0000308\sigma^2 + 1.79R^2)\,dt - 0.00542R\sigma \circ dW_A + (0.0141\frac{1}{R} - 0.0436R)\sigma \circ dW_B. \quad (6.10)$$

From simulation with the same noise intensity $\sigma = 0.54$ used in Chapter 2 a typical radius is 0.36 to 0.52 for H3p and 0.015 to 0.45 for H3u, so from the small multiplicative noise coefficients (parameters q, s_a and s_b of Eq. (6.8)) we can see that multiplicative noise effects are insignificant in this system in comparison to the additive noise, parameter p of Eq. (6.8), for both cases. (Here we use "additive noise" to describe both the constant term in the $\mathrm{d}R$ equation and the $\frac{1}{R}$ term in the $\mathrm{d}\theta$ equation, as these together give the single noise term that is independent of state when expressed in Cartesian co-ordinates.) The most significant difference between the two scenarios is the additive noise coefficient p which is 0.511 for H3p but 0.0141 for H3u, differing by a factor of 36.

Comparing the bifurcation points H1 and H2 of Chapter 2 we see a similar contrast. At Hopf point H1 the resulting simplified system is:

$$dR = (-R^3 + (0.0138\epsilon + 0.00288\sigma^2)R + 0.0905\sigma^2\frac{1}{R})\,dt \;+\; (0.425 + 0.0191R^2)\sigma \circ dW_A,$$
$$d\theta = (-65.2 - 0.0831\epsilon - 0.0106\sigma^2 + 5.71R^2)\,dt - 0.0142R\sigma \circ dW_A + (0.425\frac{1}{R} - 0.00133R)\sigma \circ dW_B. \quad (6.11)$$

Whereas at point H2, with correlated noise of the same intensity $\sigma = \sqrt{\sigma_p^2 + \sigma_u^2}$, we obtain:

$$dR = (-R^3 + (0.0166\epsilon - 0.000353\sigma^2)R + 0.00893\sigma^2\frac{1}{R})\,dt \;+\; (0.134 - 0.00211R^2)\sigma \circ dW_A,$$
$$d\theta = (-70.7 - 0.0181\epsilon - 0.000223\sigma^2 + 1.14R^2)\,dt + 0.00305R\sigma \circ dW_A + (0.134\frac{1}{R} - 0.00306R)\sigma \circ dW_B \quad (6.12)$$

As with the previous case the multiplicative noise terms are insignificant compared to the additive noise terms. (At typical radii of 0.32 to 0.50 for H1 and 0.13 to 0.44 for H2). But the coefficient of additive noise affecting the 2D oscillations is three times greater at H1 (0.425) than at H2 (0.134).

The shear parameter b also differs significantly between bifurcation points H1 and H2, and the base frequency ω also differs.

6.4 Discussion

These oscillation parameter values explain the behavior seen in Chapter 2. Multiplicative noise is not responsible. Instead, in the amplitude (dR) equations the much greater additive noise intensity of scenarios H1 and H3p increases amplitude variance, while in the angle ($d\theta$) equations the greater additive noise causes more rapid phase diffusion which destroys autocorrelation of the time series.

As the near-identity transformation of step 2 leaves additive noise unchanged, the additive noise terms in the reduced system are determined solely by the linear term of the system near bifurcation. For the Jansen-Rit system this term is a far-from-normal matrix, i.e. its eigenvectors are far from orthogonal, with consequences discussed in detail below.

The slow transitions in oscillation amplitude, too, are not caused by multiplicative noise. Instead this balloon shaped amplitude envelope is due to the weak attraction of the cubic term close to the bifurcation, causing slow mean reversion of amplitude. This may be demonstrated by comparing the processes $dx = -x\,dt + dW$ and $dx = -x^3\,dt + dW$. In the equations above the parameter γ_ϵ is small (0.0138, 0.0166 and 0.0167 respectively at points H1, H2 and H3). This means that even as the input firing rate is varied significantly, the system passes only slowly through the bifurcation in the normal form system. Thus in the Jansen-Rit system the cubic term dominates over the linear term for a wide range of parameter space, in which this amplitude behavior can be observed.

6.4.1 Predicting dependence of phase diffusion on R

As an example from the results above we can predict another qualitative difference between scenarios H3p and H3u. In the standard form Eq. (6.8) the total noise intensity affecting the phase evolution is

$$\sigma_\theta = \sigma\sqrt{(s_a R)^2 + (s_b R + \tfrac{p}{R})^2}. \tag{6.13}$$

Thus for scenario H3u,

$$\sigma_\theta \approx \sigma\sqrt{0.000198\frac{1}{R^2} + 0.00123 + 0.00193R^2},$$

whereas for H3p,

$$\sigma_\theta \approx \sigma\sqrt{0.261\frac{1}{R^2} - 0.00939 + 0.0000915R^2}.$$

As shown in Figure 6.2 for H3p the phase noise decreases monotonically as a function of R when $R < 1$, whereas for H3u phase noise will be minimal at $R = 0.566$ and will then increase again for larger oscillation amplitudes. In the Jansen-Rit model scenario H3u this corresponds to increasing the mean pyramidal input firing rate $\langle p \rangle$ beyond $100\ \mathrm{s}^{-1}$. In time series statistics increased phase noise will manifest as a faster rate of phase diffusion.

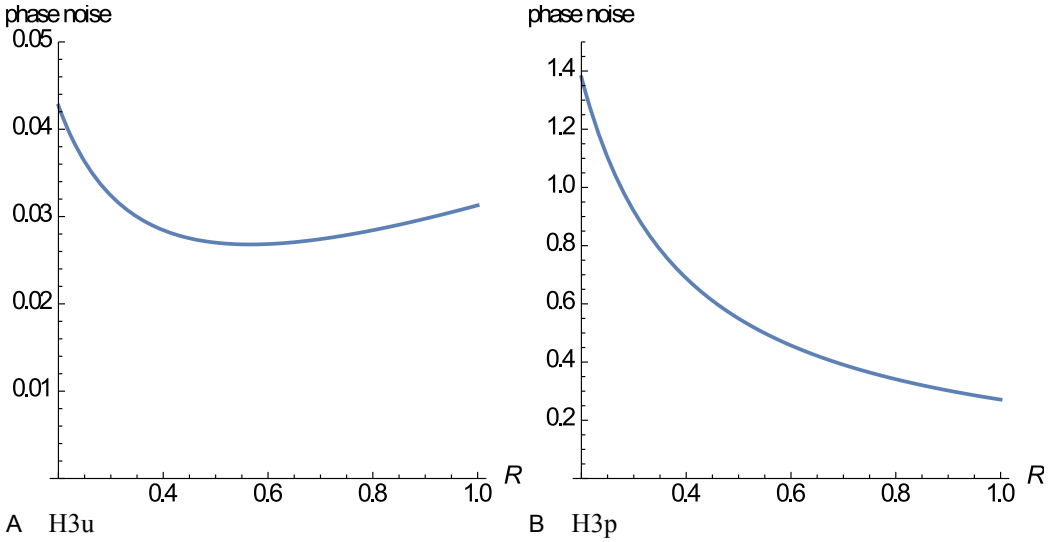


Figure 6.2: For $R \leq 1$ predicted intensity of phase noise as a function of oscillation amplitude R is non-monotonic for scenario H3u (A) but decreases monotonically for scenario H3p.

6.4.2 Effect of non-orthogonal eigenspaces

In Chapter 2 we noted it was the reference plane normal to the stable eigenspace that matters for assessing the effect of the noise on oscillations, and that noise alignment with the center eigenspace of the bifurcation was irrelevant. This reference plane is exactly the left eigenspace of the bifurcating eigenvalues. By a basic theorem of linear algebra left eigenvectors of a particular eigenvalue are normal to all right eigenvectors of the other distinct eigenvalues.

From the perspective of the normal form transformation we can see why the left eigenspaces and not right eigenspaces of the Jacobian matter for assessing the affect of noise. Viewing the system in the normal form coordinates where unperturbed oscillations lie on a circle and where the stable manifold is normal to this circle, perturbations in the plane of the circle affect oscillation phase and amplitude, while perturbations in the direction of the stable manifold do not. This feature of the normal form flow is formally described as the preservation of the stable fibration, proved as Theorem 5.1.5 in Murdock (2003).

Now for any driving noise perturbation that component of the noise in the plane of the circle is given by the inner product of the noise vector with basis directions of the center eigenspace plane. But directions in state space, like the axes, are covariant vector quantities, not contravariant like the flow vector field $\mathbf{f}(\mathbf{x})$. The latter belongs to the tangent space of the state space whereas the former are elements of the dual space. So when transforming back to the original coordinates to derive the direction sensitive to noise perturbation each unit dual-vector $\mathbf{e}$ must be transformed covariantly, that is, $\sum_i e_i M^i_j$ rather than $\sum_j M^i_j e^j$. Transforming in this way the pre-image of the normal form center eigenspace directions are then exactly the center left-eigenvectors of the Jacobian $D\mathbf{f}(\mathbf{x})$ in the original system.

In a system where the Jacobian is approximately normal, its left and right eigenspaces coincide. The directions sensitive to perturbation are then also the center eigenspace directions in the original system. But the Jansen-Rit system is unusual in that its Jacobian is extremely far from normal, leaving the left and right center eigenspaces nearly perpendicular to each other. This is why the expected effects of perturbing input to pyramidal and spiny stellate cells were reversed. The effect of non-orthogonal eigenspaces on response to perturbation is the same phenomenon that is known as “excess noise” in

laser physics (Hamel and Woerdman, 1989; Siegman, 1989; New, 1995; Firth and Yao, 2005) and "transient gain" in fluid dynamics (Grossmann, 2000; Heaton and Peake, 2007).

6.4.3 Neurobiological implications

Careful modeling of noise inputs is needed

Viewing the Jansen-Rit neural mass model in normal form coordinates has made clear why very different behavior resulted if noise perturbation was to the pyramidal cell population rather than the spiny stellate population. (Which Jansen et al. (1993) suggested should not make a difference). For biophysical models to be comparable to experimental time series statistics, the modeling of stochasticity requires the same attention as the modeling of the deterministic flow, since the noise model can have significant effects on the resulting statistics that are compared to experiment.

Implications for autocorrelation in neuroimaging time series

Long autocorrelation times may or may not manifest when linear stability is lost. As seen above it can depend on the orientation of the noise in the state space and more generally on the stability of oscillation frequency. Therefore as an indicator of the loss of linear stability we suggest using the autocorrelation of the Hilbert amplitude of the measured signal, rather than the autocorrelation of the raw signal.

New possible mechanism for amplitude fluctuations in alpha/beta oscillations

If brain oscillations operate near the point of bifurcation, which would permit the presence of oscillations to be changed dynamically to facilitate dynamic synchrony between brain regions, then slow relaxation times in amplitude may apply due to the weaker $-x^3$ attraction near a Hopf bifurcation discussed above. In the real brain whether this occurred only close to the transition point or beyond it would depend on whether the true value of parameter γ_ϵ is small, as it is in the Jansen-Rit model considered above.

This effect would manifest in neuroimaging time series as the balloon-shaped variation of oscillation amplitude envelope on slow time scales similar to those seen in the simulations of Chapter 2. This gives another possible mechanism for the slow time scale amplitude changes observed in brain oscillations

at different frequencies, consistent with the idea that operating near bifurcation may be adaptive for brain function.

Chapter 7

Review and conclusions

7.1 New insights

In an attempt to understand the interesting effects of noise near a Hopf bifurcation observed in our simulations, we have developed a new method which may be broadly useful beyond this original setting.

The ongoing investigation into whether the brain operates near the point of a multi-scale continuous phase transition is an important one. So it is worth noting that we have confirmed that phenomena associated with criticality at a phase transition, such as power law scaling over several orders of magnitude for critical fluctuation duration and power, are also present near a low-dimensional Hopf bifurcation in the presence of noise without a phase transition. This may exemplify the analogy between macroscopic bifurcations in fluctuating systems and non-equilibrium phase transitions suggested by Haken (1978).

Regarding statistical indicators of instability and bifurcation in time series, the sensitive dependence of these indicators on both the local flow and the specific noise perturbations has made clear that univariate time series are best interpreted in the context of a model, rather than a data-driven approach that assumes that a signature for detecting each type of transition is generic. The methods developed in this thesis provide one way to use such a model.

7.2 Limitations

We used the deterministic normal form transformation together with asymptotic approximations which limit the applicability to either weak noise or strong stability of the center manifold.

The current method only applies to transitions that are local bifurcations of equilibria, or to bistability in the case of the Generalized Hopf normal form, though an analogous approach may be possible to simplify the flow near an arbitrary limit cycle to treat bifurcations of limit cycles, following setting (3) of Murdock (2003).

The simple standard form of Section 4.5 results from mapping to a weak approximation of the stochastic process. The resulting weak model always has two driving Wiener processes, regardless of whether the original system had one or many. This is actually advantageous in the case of multiple driving Wiener processes, but it is possible that more advanced methods may allow reduction to a similar level of simplicity in a strong approximation, so that sample paths correspond. This would be an improvement particularly in the case of the original stochastic system having only one driving Wiener process.

Finally, mappings from a detailed biophysical model will only be of practical use for clinical time series if the biophysical model itself is accurate in representing the oscillatory behavior. A comparison of current neural mass models to empirical power spectra and coherence shows that our field is still progressing towards that point. By connecting models to univariate time series dynamics it is hoped this will contribute to evolving increasingly accurate models, at which time the methods of this thesis will have direct clinical application.

7.3 Future work

We have connected stochastic biophysical models to a simple standard form. To complete the goal of interpreting clinical time series the remaining task is to derive the quantitative relationships between the oscillation terms in the standard form and the oscillation statistics, in particular autocorrelation of oscillation amplitude.

The transformation and approximation methods presented in the thesis are not limited to Hopf bifurcations but can be applied to all other local bifurcation types. This will have useful application to neuroscience models where a transition to oscillation occurs through a saddle-node on invariant circle (SNIC) bifurcation, even though the bifurcation is technically a global one. In this case stochastic fluctuations make a crucial difference to dynamics in the slow region, specifically the dependence of the distribution of oscillator period on system and noise parameters, and this relationship can potentially be determined from a biophysical model by the stochastic methods of this thesis applied to the saddle-node bifurcation of the SNIC.

It remains to characterize rigorously the class of oscillator statistics that are preserved by the Poincaré normal form transformation near a Hopf bifurcation. We proved only that the distribution of periods is preserved, and conjectured that several other statistics are approximately preserved on the basis of large scale simulations.

In the 8-dimensional Jansen-Rit model the interesting behavior arose ultimately from the fact that the left and right center eigenspaces were close to orthogonal. We suggest that this becomes increasingly likely (for various random matrix Jacobians) as the dimension d of the system is increased. We have started some initial work to investigate this with Monte Carlo simulations in d dimensions.

Implementing the more sophisticated normal form transformation algorithms based on Lie theory (Murdock, 2003) will provide some further benefits: Firstly, the algorithms are more computationally efficient. Secondly, they explicitly find the inverse transformation as well as the forward transformation, which has important applications. For example, having the inverse transformation will then allow a phase response curve of the Poincaré normal form to be transformed back to determine phase sensitivity as a function of the original biological coordinates (Brown et al. 2004).

Appendix A

Derivations for the deterministic Jansen-Rit system

The deterministic Jansen-Rit model, with input firing rates $p(t)$ and $u(t)$ to pyramidal and spiny stellate populations respectively, is given as a system of first order ordinary differential equations in Eq. (3.37).

A.1 Location of bifurcation points

In this system it is possible to solve for the deterministic equilibrium point $\mathbf{v}_c = (v_{c1}, v_{c2}, v_{c3}, v_{c4}, v_{c5}, v_{c6}, v_{c7}, v_{c8},)^T$ and bifurcation point p_c directly: from Eq. (3.37) with $\mathbf{f}(\mathbf{v}_c) = \mathbf{0}$,

$$\begin{aligned}
S^{-1}(\phi) &= \rho_2 - \ln(2e_0/\phi - 1)/\rho_1, \\
v_{c5} &= v_{c6} = v_{c7} = v_{c8} = 0, \\
v_{c4}(v_{c3}) &= S^{-1}(\kappa_i v_{c3}/(\gamma_4 H_i)), \\
v_{c2}(v_{c3}) &= S^{-1}(\kappa_e v_{c4}(v_{c3})/(\gamma_3 H_e)) + v_{c3}, \\
v_{c1}(v_{c3}, u) &= H_e(\gamma_1 \kappa_e v_{c4}(v_{c3})/(\gamma_3 H_e) + u)/\kappa_e, \\
p_c(v_{c3}, u) &= \kappa_e v_{c2}(v_{c3})/H_e - \gamma_2 S(v_{c1}(v_{c3}, u)),
\end{aligned} \tag{A.1}$$

thus obtaining p_c and $\mathbf{v}_c$ as explicit functions of v_{c3} and u. We then have an explicit expression for the Jacobian matrix A at bifurcation, as a function of v_{c3} and u:

$$A(v_{c3}, u) = D\mathbf{f}(\mathbf{v})|_{\mathbf{v}=\mathbf{v}_c(v_{c3},u)}. \tag{A.2}$$

The roots of the characteristic polynomial of $A(v_{c3}, u)$ then give the bifurcating eigenvalues $\alpha + i\beta$ as functions of v_{c3} and u, whence $\alpha(v_{c3}, u) = 0$ defines the bifurcation points exactly and can be solved for v_{c3} to arbitrary precision given u. Eq. (A.1) then gives p_c and $\mathbf{v}_c$. For supercritical Hopf points H1 and H2 we thus obtain

$$\begin{aligned}
&\text{H1:} \quad p_c = 89.829\,\text{s}^{-1}, \quad \mathbf{v}_c = (13.208, 20.165, 13.425, 3.3021, 0, 0, 0, 0)^{\text{T}}\ \text{mV}, \\
&\text{H2:} \quad p_c = 73.009\,\text{s}^{-1}, \quad \mathbf{v}_c = (21.852, 19.920, 13.225, 3.2694, 0, 0, 0, 0)^{\text{T}}\ \text{mV},
\end{aligned} \tag{A.3}$$

respectively. To locate point H3 (defined in Chapter 2 as the point where the Hopf branch intersects the line $\langle u \rangle = \langle p \rangle$), since $p_c(v_{c3}(u), u)$ is a continuous, monotonic function of u near this point, the required value of u can be computed iteratively to arbitrary precision as the limit of the sequence $u_{n+1} = \frac{1}{2}(u_n + p_c(v_{c3}(u_n), u_n)$. This gives

$$\text{H3:} \quad p_c = u = 80.346\,\text{s}^{-1}, \quad \mathbf{v}_c = (15.779, 20.088, 13.362, 3.2918, 0, 0, 0, 0)^{\text{T}}\ \text{mV}. \tag{A.4}$$

A.2 Normal form transformation

Initial linear transformation:

At point H1, the Jacobian $A = D\mathbf{f}(\mathbf{v})|_{\mathbf{v}_c}$ is approximately

$$A = \begin{bmatrix}
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
-10000 & 29432.3 & -29432.3 & 0 & -200 & 0 & 0 & 0 \\
1675.66 & -10000 & 0 & 0 & 0 & -200 & 0 & 0 \\
0 & 0 & -2500 & 15397 & 0 & 0 & -100 & 0 \\
0 & 7358.06 & -7358.06 & -10000 & 0 & 0 & 0 & -200
\end{bmatrix}. \tag{A.5}$$

This matrix has condition number (ratio of largest to smallest singular values) of 44822. Rebalancing (Parlett and Reinsch, 1969) with diagonal matrix $T = \text{diag}(\frac{1}{16}, \frac{1}{32}, \frac{1}{32}, \frac{1}{32}, 8, 4, 4, 4)$ reduces the condition number of $T^{-1}AT$ to 17.96. Computing Jordan decomposition $T^{-1}AT = GBG^{-1}$ so that the

new Jacobian $B = G^{-1}T^{-1}A\,T\,G$ is in Jordan real form, the overall linear transformation M to be used in step 1 is

$$M = TG = \begin{bmatrix} 0 & -0.49079 & -0.180728 & -0.110331 & 0 & 0 & -0.0988392 & -0.199643 \\ -0.0528038 & -0.0232788 & 0.0497681 & -0.00789791 & -0.246353 & -0.0123176 & -0.043254 & -0.848442 \\ -0.270252 & 0.0725844 & 0.0339665 & -0.0560267 & -0.246353 & -0.0123176 & -0.0303954 & -0.845768 \\ 0 & -0.122698 & -0.045182 & -0.0275829 & -0.04 & -0.0004 & -0.0247098 & -0.0499106 \\ 32 & 0 & 32 & 0 & 0 & 0 & 16 & 16 \\ 1.5178 & -3.44286 & -5.7976 & 4.93775 & 24.6353 & 0.985411 & 7.00193 & 67.9969 \\ -4.73257 & -17.6207 & 0.0305537 & 9.90153 & 24.6353 & 0.985411 & 4.92038 & 67.7825 \\ 8 & 0 & 8 & 0 & 4 & 0 & 4 & 4 \end{bmatrix}, \tag{A.6}$$

with new Jacobian $$B = \begin{bmatrix} 0 & 65.2010 & 0 & 0 & 0 & 0 & 0 & 0 \\ -65.2010 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -128.989 & 78.7454 & 0 & 0 & 0 & 0 \\ 0 & 0 & -78.7454 & -128.989 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -100 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -100 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -161.879 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -80.1432 \end{bmatrix}. \tag{A.7}$$

After the translation and linear transformation $\mathbf{v} - \mathbf{v}_c = M\mathbf{x}$ is applied to Eq. (6.3) the linearly transformed system is then $\dot{\mathbf{x}} = M^{-1}\mathbf{f}(M\mathbf{x} + \mathbf{v}_c)$, that is,

$$\begin{aligned}
\dot{x_1} &= (-1919.3 + 1/(0.0018319 + 0.0012107\exp(-0.12177x_1 + 0.053683x_2 - 0.0088489x_3 - 0.026952x_4 \\
&\quad + 0.0072008x_7 + 0.0014977x_8)) + 1/(-0.00023261 - 0.0010539\exp(0.068711x_2 + 0.025302x_3 + 0.015446x_4 \\
&\quad + 0.0224x_5 + 0.000224x_6 + 0.013837x_7 + 0.02795x_8)) + 1/(0.00041497 + 7.3273{\times}10^{-6}\exp(0.27484x_2 \\
&\quad + 0.10121x_3 + 0.061786x_4 + 0.05535x_7 + 0.1118x_8)) - 15.926x_1 + 39.762x_2 - 13.11x_3 - 10.822x_4 - 14.264x_5 \\
&\quad - 0.14264x_6 - 5.5954x_7 - 13.008x_8 + 4.4626\epsilon) \\
\dot{x_2} &= (1475.9 + 1/(-0.00073893 - 0.00048835\exp(-0.12177x_1 + 0.053683x_2 - 0.0088489x_3 - 0.026952x_4 \\
&\quad + 0.0072008x_7 + 0.0014977x_8)) + 1/(-0.00030867 - 0.0013984\exp(0.068711x_2 + 0.025302x_3 + 0.015446x_4 \\
&\quad + 0.0224x_5 + 0.000224x_6 + 0.013837x_7 + 0.02795x_8)) + 1/(-0.013041 - 0.00023027\exp(0.27484x_2 \\
&\quad + 0.10121x_3 + 0.061786x_4 + 0.05535x_7 + 0.1118x_8)) - 25.72x_1 - 50.737x_2 - 9.4049x_3 + 1.2456x_4 - 10.749x_5 \\
&\quad - 0.10749x_6 - 9.0472x_7 - 14.044x_8 - 0.142\epsilon) \\
\dot{x_3} &= (14804. + 1/(0.00029395 + 0.00019427\exp(-0.12177x_1 + 0.053683x_2 - 0.0088489x_3 - 0.026952x_4 \\
&\quad + 0.0072008x_7 + 0.0014977x_8)) + 1/(0.0001062 + 0.00048115\exp(0.068711x_2 + 0.025302x_3 + 0.015446x_4 \\
&\quad + 0.0224x_5 + 0.000224x_6 + 0.013837x_7 + 0.02795x_8)) + 1/(-0.00005296 - 9.3514{\times}10^{-7}\exp(0.27484x_2 \\
&\quad + 0.10121x_3 + 0.061786x_4 + 0.05535x_7 + 0.1118x_8)) - 99.249x_1 + 51.105x_2 - 133.49x_3 + 58.431x_4 \\
&\quad + 31.242x_5 + 0.31242x_6 + 7.3493x_7 + 4.2109x_8 - 34.967\epsilon) \\
\dot{x_4} &= (-2943. + 1/(-0.00021213 - 0.0001402\exp(-0.12177x_1 + 0.053683x_2 - 0.0088489x_3 - 0.026952x_4 \\
&\quad + 0.0072008x_7 + 0.0014977x_8)) + 1/(0.00014621 + 0.0006624\exp(0.068711x_2 + 0.025302x_3 + 0.015446x_4 \\
&\quad + 0.0224x_5 + 0.000224x_6 + 0.013837x_7 + 0.02795x_8)) + 1/(0.00021622 + 3.8179{\times}10^{-6}\exp(0.27484x_2 \\
&\quad + 0.10121x_3 + 0.061786x_4 + 0.05535x_7 + 0.1118x_8)) + 137.53x_1 + 30.652x_2 - 35.138x_3 - 78.029x_4 \\
&\quad + 22.693x_5 + 0.22693x_6 + 10.25x_7 + 35.44x_8 + 8.5645\epsilon) \\
\dot{x_5} &= (-100x_5 + x_6) \\
\dot{x_6} &= -100x_6 \\
\dot{x_7} &= (-36878. + 1/(0.00017042 + 0.00011263\exp(-0.12177x_1 + 0.053683x_2 - 0.0088489x_3 - 0.026952x_4 \\
&\quad + 0.0072008x_7 + 0.0014977x_8)) + 1/(-0.000085655 - 0.00038807\exp(0.068711x_2 + 0.025302x_3 \\
&\quad + 0.015446x_4 + 0.0224x_5 + 0.000224x_6 + 0.013837x_7 + 0.02795x_8)) + 1/(0.000027714 \\
&\quad + 4.8936{\times}10^{-7}\exp(0.27484x_2 + 0.10121x_3 + 0.061786x_4 + 0.05535x_7 + 0.1118x_8)) - 171.18x_1 + 125.73x_2 \\
&\quad + 6.07x_3 - 26.589x_4 - 38.735x_5 - 0.38735x_6 - 141.63x_7 + 22.552x_8 + 66.819\epsilon) \\
\dot{x_8} &= (2854.8 + 1/(-0.01903 - 0.012577\exp(-0.12177x_1 + 0.053683x_2 - 0.0088489x_3 - 0.026952x_4 + 0.0072008x_7
\end{aligned}$$

$$+ 0.0014977x_8)) + 1/(0.0006943 + 0.0031456\exp(0.068711x_2 + 0.025302x_3 + 0.015446x_4 + 0.0224x_5 + 0.000224x_6 + 0.013837x_7 + 0.02795x_8)) + 1/(-0.00031867 - 5.6269\times10^{-6}\exp(0.27484x_2 + 0.10121x_3 + 0.061786x_4 + 0.05535x_7 + 0.1118x_8)) + 1.533x_1 - 0.7222x_2 + 0.094334x_3 + 0.32889x_4 + 4.7787x_5 + 0.047787x_6 - 0.09999x_7 - 80.181x_8 - 5.8111\epsilon). \quad \text{(A.8)}$$

Near-identity transformation:

In preparation the right-hand-side of Eq. (A.8) is now approximated near the origin in phase space and parameter space by a Taylor series truncated to third order. As the Jansen-Rit Jacobian B is not diagonalizable over $\mathbb{C}$, we use the inner product normal form style discussed in Section 3.3.2. We form a 9-dimensional extended system by promoting the bifurcation parameter ϵ to a dynamical variable, introducing $\alpha = \sqrt{\epsilon}$ so that $O(\alpha) = O(x_i)$ near bifurcation, with equation of motion $\dot{\alpha} = 0$ so this extended system will have a 3-dimensional center manifold. The normal form transformation $\mathbf{x} = \mathbf{y} + \mathbf{q}(\mathbf{y}, \alpha)$ is then computed to third order by the iterative method of Section 3.3.2 giving this α-dependent change of coordinates:

$$\begin{aligned}
x_1 = {} & y_1 - 0.002530y_1^2 - 0.0007372y_1^3 - 0.02128y_1y_2 + 0.000455y_1^2y_2 + 0.00116y_2^2 - 0.001604y_1y_2^2 - 0.00003821y_2^3 \\
& - 0.00199y_1y_3 + 0.0002932y_1^2y_3 + 0.006288y_2y_3 - 0.000484y_1y_2y_3 - 0.0013y_2^2y_3 + 0.0007812y_3^2 - 0.0001419y_1y_3^2 \\
& - 0.0002229y_2y_3^2 + 0.00001517y_3^3 - 0.001713y_1y_4 - 0.00006445y_1^2y_4 + 0.008011y_2y_4 + 0.001762y_1y_2y_4 \\
& + 0.0001973y_2^2y_4 + 0.001403y_3y_4 + 0.0002499y_1y_3y_4 - 0.0002627y_2y_3y_4 + 0.00002713y_3^2y_4 + 0.0008333y_4^2 \\
& + 0.000395y_1y_4^2 + 0.00003935y_2y_4^2 + 0.000026y_3y_4^2 + 0.00001235y_4^3 - 0.00002612y_1y_5 + 0.0001229y_1^2y_5 \\
& + 0.003148y_2y_5 - 0.0003958y_1y_2y_5 + 0.0002887y_2^2y_5 + 0.0006217y_3y_5 - 0.0000877y_1y_3y_5 - 0.0001041y_2y_3y_5 \\
& - 1.762\times10^{-6}y_3^2y_5 + 0.0005858y_4y_5 + 0.0001152y_1y_4y_5 + 0.00002166y_2y_4y_5 - 3.496\times10^{-6}y_3y_4y_5 \\
& + 2.213\times10^{-6}y_4^2y_5 + 0.0003477y_5^2 - 0.00002571y_1y_5^2 + 0.000044y_2y_5^2 + 2.85\times10^{-6}y_3y_5^2 + 3.987\times10^{-6}y_4y_5^2 \\
& + 1.581\times10^{-6}y_5^3 + 0.00001439y_1y_6 - 3.104\times10^{-7}y_1^2y_6 + 0.00005133y_2y_6 + 3.091\times10^{-6}y_1y_2y_6 + 6.329\times10^{-7}y_2^2y_6 \\
& + 7.861\times10^{-6}y_3y_6 + 4.628\times10^{-7}y_1y_3y_6 - 2.408\times10^{-6}y_2y_3y_6 - 9.131\times10^{-8}y_3^2y_6 + 7.621\times10^{-6}y_4y_6 \\
& + 3.038\times10^{-6}y_1y_4y_6 - 2.758\times10^{-7}y_2y_4y_6 - 1.007\times10^{-7}y_3y_4y_6 + 2.315\times10^{-8}y_4^2y_6 + 8.63\times10^{-6}y_5y_6 \\
& - 5.801\times10^{-7}y_1y_5y_6 + 2.481\times10^{-6}y_2y_5y_6 + 3.945\times10^{-7}y_3y_5y_6 + 4.257\times10^{-7}y_4y_5y_6 + 1.99\times10^{-7}y_5^2y_6 \\
& + 5.097\times10^{-8}y_6^2 - 1.675\times10^{-8}y_1y_6^2 + 6.849\times10^{-8}y_2y_6^2 + 1.359\times10^{-8}y_3y_6^2 + 1.332\times10^{-8}y_4y_6^2 + 7.871\times10^{-9}y_5y_6^2 \\
& - 0.001413y_1y_7 + 0.0002356y_1^2y_7 + 0.004485y_2y_7 - 0.0001436y_1y_2y_7 - 0.0005201y_2^2y_7 + 0.0009191y_3y_7 \\
& - 0.0001503y_1y_3y_7 - 0.0002847y_2y_3y_7 + 0.00001962y_3^2y_7 + 0.0008702y_4y_7 + 0.0001841y_1y_4y_7 - 0.0001075y_2y_4y_7 \\
& + 0.00002927y_3y_4y_7 + 0.0000181y_4^2y_7 + 0.000368y_5y_7 - 0.00005045y_1y_5y_7 - 0.00004089y_2y_5y_7 \\
& - 2.213\times10^{-6}y_3y_5y_7 + 1.564\times10^{-6}y_4y_5y_7 + 2.829\times10^{-6}y_5^2y_7 + 4.597\times10^{-6}y_6y_7 + 2.319\times10^{-7}y_1y_6y_7 \\
& - 8.693\times10^{-7}y_2y_6y_7 - 4.806\times10^{-8}y_3y_6y_7 + 4.053\times10^{-8}y_4y_6y_7 + 2.938\times10^{-7}y_5y_6y_7 + 8.813\times10^{-9}y_6^2y_7 \\
& + 0.0002826y_7^2 - 0.00004164y_1y_7^2 - 0.00007229y_2y_7^2 + 0.0000102y_3y_7^2 + 0.00001081y_4y_7^2 + 5.042\times10^{-7}y_5y_7^2 \\
& + 1.926\times10^{-8}y_6y_7^2 + 2.358\times10^{-6}y_7^3 - 0.003178y_1y_8 + 0.0009023y_1^2y_8 + 0.01433y_2y_8 - 0.002151y_1y_2y_8 \\
& + 0.0006395y_2^2y_8 + 0.002276y_3y_8 - 0.0004707y_1y_3y_8 - 0.0004133y_2y_3y_8 + 0.00001552y_3^2y_8 + 0.002454y_4y_8 \\
& + 0.0002549y_1y_4y_8 + 0.0001709y_2y_4y_8 + 0.00004251y_3y_4y_8 + 0.00004155y_4^2y_8 + 0.0009084y_5y_8 \\
& - 0.0002351y_1y_5y_8 + 0.0002266y_2y_5y_8 + 0.00001147y_3y_5y_8 + 0.00002897y_4y_5y_8 + 0.00001851y_5^2y_8 \\
& + 0.00001107y_6y_8 - 3.537\times10^{-6}y_1y_6y_8 + 0.00001116y_2y_6y_8 + 1.952\times10^{-6}y_3y_6y_8 + 2.165\times10^{-6}y_4y_6y_8 \\
& + 1.768\times10^{-6}y_5y_6y_8 + 3.343\times10^{-8}y_6^2y_8 + 0.001438y_7y_8 - 0.000302y_1y_7y_8 - 0.00004886y_2y_7y_8 \\
& + 0.00003913y_3y_7y_8 + 0.00005972y_4y_7y_8 + 0.00002395y_5y_7y_8 + 1.504\times10^{-6}y_6y_7y_8 + 0.00002302y_7^2y_8 \\
& + 0.001934y_8^2 + 0.003221y_1y_8^2 - 0.01174y_2y_8^2 - 0.002437y_3y_8^2 - 0.002263y_4y_8^2 - 0.0009338y_5y_8^2 \\
& - 8.683\times10^{-6}y_6y_8^2 - 0.00145y_7y_8^2 - 0.003802y_8^3 - 0.002178\alpha^2 + 0.001721y_1\alpha^2 + 0.0003924y_2\alpha^2 - 0.0002968y_3\alpha^2 \\
& - 0.0003115y_4\alpha^2 - 0.00008638y_5\alpha^2 - 3.778\times10^{-7}y_6\alpha^2 - 0.0001884y_7\alpha^2 - 0.0004541y_8\alpha^2
\end{aligned}$$

$$\begin{aligned}
x_2 = {} & y_2 + 0.02432y_1^2 - 0.0005537y_1^3 - 0.01006y_1y_2 + 0.0004057y_1^2y_2 + 0.017y_2^2 + 0.00016y_1y_3 + 0.0005278y_1^2y_3 \\
& + 0.00346y_2y_3 + 0.0002419y_1y_2y_3 - 0.001208y_2^2y_3 + 0.0003411y_3^2 - 0.00006238y_1y_3^2 - 0.000145y_2y_3^2
\end{aligned}$$

$$
\begin{aligned}
&+ 2.555\times10^{-6}y_3^3 - 0.005681y_1y_4 - 0.0007362y_1^2y_4 + 0.006387y_2y_4 + 0.001571y_1y_2y_4 - 0.0001034y_2^2y_4\\
&+ 0.0006598y_3y_4 + 0.00009545y_1y_3y_4 - 0.0002642y_2y_3y_4 + 3.504\times10^{-6}y_3^2y_4 + 0.0004511y_4^2 + 0.0002183y_1y_4^2\\
&- 0.00002776y_2y_4^2 + 3.666\times10^{-6}y_3y_4^2 + 4.469\times10^{-6}y_4^3 - 0.003108y_1y_5 + 0.0003779y_1^2y_5 + 0.004741y_2y_5\\
&- 0.000288y_1y_2y_5 + 0.0002507y_2^2y_5 + 0.0006427y_3y_5 - 0.00009049y_1y_3y_5 - 0.0001826y_2y_3y_5 + 3.333\times10^{-8}y_3^2y_5\\
&+ 0.0008507y_4y_5 + 0.0001916y_1y_4y_5 - 0.00003535y_2y_4y_5 - 3.285\times10^{-6}y_3y_4y_5 + 2.583\times10^{-6}y_4^2y_5 + 0.0004976y_5^2\\
&- 0.00003073y_1y_5^2 + 0.00003662y_2y_5^2 + 1.151\times10^{-8}y_3y_5^2 + 2.566\times10^{-6}y_4y_5^2 + 1.836\times10^{-6}y_5^3 - 0.0000738y_1y_6\\
&+ 2.282\times10^{-6}y_1^2y_6 + 0.00007991y_2y_6 + 4.35\times10^{-7}y_1y_2y_6 + 3.149\times10^{-6}y_2^2y_6 + 8.058\times10^{-6}y_3y_6\\
&- 6.436\times10^{-7}y_1y_3y_6 - 3.27\times10^{-6}y_2y_3y_6 - 2.781\times10^{-8}y_3^2y_6 + 0.00001328y_4y_6 + 3.724\times10^{-6}y_1y_4y_6\\
&- 8.802\times10^{-7}y_2y_4y_6 - 7.706\times10^{-8}y_3y_4y_6 + 3.568\times10^{-8}y_4^2y_6 + 0.00001547y_5y_6 - 5.159\times10^{-7}y_1y_5y_6\\
&+ 1.712\times10^{-6}y_2y_5y_6 + 1.237\times10^{-7}y_3y_5y_6 + 2.686\times10^{-7}y_4y_5y_6 + 2.387\times10^{-7}y_5^2y_6 + 1.324\times10^{-7}y_6^2\\
&- 1.316\times10^{-8}y_1y_6^2 + 4.401\times10^{-8}y_2y_6^2 + 5.59\times10^{-9}y_3y_6^2 + 8.41\times10^{-9}y_4y_6^2 + 9.847\times10^{-9}y_5y_6^2 - 0.001335y_1y_7\\
&- 0.00005327y_1^2y_7 + 0.002791y_2y_7 + 0.0002029y_1y_2y_7 - 0.000487y_2^2y_7 + 0.000413y_3y_7 - 0.0000712y_1y_3y_7\\
&- 0.0001964y_2y_3y_7 + 1.612\times10^{-6}y_3^2y_7 + 0.0005613y_4y_7 + 0.0001156y_1y_4y_7 - 0.0001168y_2y_4y_7\\
&+ 2.487\times10^{-6}y_3y_4y_7 + 5.902\times10^{-6}y_4^2y_7 + 0.0004542y_5y_7 - 0.00003057y_1y_5y_7 - 0.00008036y_2y_5y_7\\
&- 1.975\times10^{-6}y_3y_5y_7 + 5.92\times10^{-7}y_4y_5y_7 + 1.08\times10^{-6}y_5^2y_7 + 6.505\times10^{-6}y_6y_7 + 2.102\times10^{-7}y_1y_6y_7\\
&- 1.272\times10^{-6}y_2y_6y_7 - 3.506\times10^{-8}y_3y_6y_7 + 1.702\times10^{-8}y_4y_6y_7 + 1.366\times10^{-7}y_5y_6y_7 + 4.244\times10^{-9}y_6^2y_7\\
&+ 0.0001362y_7^2 - 8.513\times10^{-6}y_1y_7^2 - 0.00005474y_2y_7^2 + 4.208\times10^{-7}y_3y_7^2 + 1.369\times10^{-6}y_4y_7^2 + 6.518\times10^{-8}y_5y_7^2\\
&+ 6.046\times10^{-9}y_6y_7^2 + 2.375\times10^{-7}y_7^3 - 0.0105y_1y_8 + 0.001537y_1^2y_8 + 0.009908y_2y_8 - 0.0006049y_1y_2y_8\\
&+ 0.0001646y_2^2y_8 + 0.001021y_3y_8 - 0.000267y_1y_3y_8 - 0.0004359y_2y_3y_8 - 8.214\times10^{-6}y_3^2y_8 + 0.001753y_4y_8\\
&+ 0.0002523y_1y_4y_8 - 0.0001227y_2y_4y_8 - 0.00001105y_3y_4y_8 + 0.00001589y_4^2y_8 + 0.001393y_5y_8 - 0.0002351y_1y_5y_8\\
&+ 0.00007941y_2y_5y_8 - 5.562\times10^{-6}y_3y_5y_8 + 0.0000162y_4y_5y_8 + 0.00001726y_5^2y_8 + 0.00002239y_6y_8\\
&- 4.504\times10^{-6}y_1y_6y_8 + 6.078\times10^{-6}y_2y_6y_8 + 7.189\times10^{-7}y_3y_6y_8 + 1.336\times10^{-6}y_4y_6y_8 + 1.696\times10^{-6}y_5y_6y_8\\
&+ 2.843\times10^{-8}y_6^2y_8 + 0.0008417y_7y_8 - 0.0001216y_1y_7y_8 - 0.0001374y_2y_7y_8 - 6.964\times10^{-7}y_3y_7y_8\\
&+ 0.00001382y_4y_7y_8 + 0.000013y_5y_7y_8 + 7.451\times10^{-7}y_6y_7y_8 + 4.458\times10^{-6}y_7^2y_8 + 0.001483y_8^2 + 0.003305y_1y_8^2\\
&- 0.007539y_2y_8^2 - 0.00112y_3y_8^2 - 0.001487y_4y_8^2 - 0.001189y_5y_8^2 - 0.00001537y_6y_8^2 - 0.0007148y_7y_8^2 - 0.002247y_8^3\\
&- 0.06844\alpha^2 - 0.00009718y_3\alpha^2 - 0.0003578y_4\alpha^2 - 0.0002387y_5\alpha^2 - 4.457\times10^{-6}y_6\alpha^2 - 0.0001302y_7\alpha^2\\
&- 0.0006302y_8\alpha^2
\end{aligned}
$$

$$
\begin{aligned}
x_3 = y_3 &+ 0.003911y_1^2 + 0.0002712y_1^3 + 0.04354y_1y_2 + 0.003925y_1^2y_2 + 0.06832y_2^2 + 0.001856y_1y_2^2 + 0.007315y_2^3\\
&+ 0.05264y_1y_3 - 0.003647y_1^2y_3 + 0.01463y_2y_3 + 0.02729y_1y_2y_3 - 0.001681y_2^2y_3 - 0.005696y_3^2 + 0.001454y_1y_3^2\\
&+ 0.001654y_2y_3^2 - 0.0001454y_3^3 + 0.0531y_1y_4 + 0.01091y_1^2y_4 - 0.03733y_2y_4 + 0.003773y_1y_2y_4 + 0.03061y_2^2y_4\\
&- 0.01284y_3y_4 - 0.001537y_1y_3y_4 + 0.002616y_2y_3y_4 - 0.0003067y_3^2y_4 - 0.009019y_4^2 - 0.00362y_1y_4^2\\
&- 0.0009328y_2y_4^2 - 0.000297y_3y_4^2 - 0.0001644y_4^3 + 0.008933y_1y_5 - 0.001089y_1^2y_5 + 0.03023y_2y_5\\
&+ 0.002246y_1y_2y_5 + 0.002592y_2^2y_5 - 0.002519y_3y_5 + 0.0004502y_1y_3y_5 - 0.00003444y_2y_3y_5 + 0.0000184y_3^2y_5\\
&- 0.001312y_4y_5 + 0.0002854y_1y_4y_5 - 0.0005966y_2y_4y_5 + 0.0000207y_3y_4y_5 - 0.00001757y_4^2y_5 - 0.0002744y_5^2\\
&+ 0.00003333y_1y_5^2 - 0.0003826y_2y_5^2 - 0.00003373y_3y_5^2 - 0.00004816y_4y_5^2 - 4.178\times10^{-6}y_5^3 + 0.0002572y_1y_6\\
&+ 0.000052y_1^2y_6 - 0.0004796y_2y_6 - 0.00004784y_1y_2y_6 - 2.237\times10^{-7}y_2^2y_6 - 0.00005112y_3y_6 - 9.739\times10^{-6}y_1y_3y_6\\
&+ 5.104\times10^{-6}y_2y_3y_6 + 7.876\times10^{-7}y_3^2y_6 - 9.728\times10^{-6}y_4y_6 - 9.437\times10^{-6}y_1y_4y_6 - 4.888\times10^{-6}y_2y_4y_6\\
&+ 9.15\times10^{-7}y_3y_4y_6 - 2.113\times10^{-7}y_4^2y_6 + 0.00002731y_5y_6 - 7.609\times10^{-7}y_1y_5y_6 - 0.00003398y_2y_5y_6\\
&- 3.658\times10^{-6}y_3y_5y_6 - 4.832\times10^{-6}y_4y_5y_6 - 5.067\times10^{-7}y_5^2y_6 + 7.637\times10^{-7}y_6^2 - 3.177\times10^{-8}y_1y_6^2\\
&- 1.091\times10^{-6}y_2y_6^2 - 1.168\times10^{-7}y_3y_6^2 - 1.493\times10^{-7}y_4y_6^2 - 1.868\times10^{-8}y_5y_6^2 - 1.285\times10^{-1}0y_6^3 - 0.009502y_1y_7\\
&- 0.00304y_1^2y_7 - 0.08187y_2y_7 - 0.001779y_1y_2y_7 + 0.0048y_2^2y_7 - 0.007743y_3y_7 + 0.001486y_1y_3y_7\\
&+ 0.002232y_2y_3y_7 - 0.0002026y_3^2y_7 - 0.009775y_4y_7 - 0.001547y_1y_4y_7 + 0.0005854y_2y_4y_7 - 0.0003324y_3y_4y_7\\
&- 0.0002372y_4^2y_7 - 0.0008755y_5y_7 + 0.0004192y_1y_5y_7 - 0.0001769y_2y_5y_7 + 0.0000105y_3y_5y_7 - 0.0000149y_4y_5y_7\\
&- 0.00002455y_5^2y_7 - 6.94\times10^{-6}y_6y_7 - 2.969\times10^{-6}y_1y_6y_7 - 7.946\times10^{-7}y_2y_6y_7 + 3.274\times10^{-7}y_3y_6y_7\\
&- 3.753\times10^{-7}y_4y_6y_7 - 2.533\times10^{-6}y_5y_6y_7 - 7.82\times10^{-8}y_6^2y_7 - 0.002545y_7^2 + 0.0003205y_1y_7^2 + 0.0005389y_2y_7^2\\
&- 0.0001087y_3y_7^2 - 0.0001206y_4y_7^2 - 4.207\times10^{-6}y_5y_7^2 - 1.57\times10^{-7}y_6y_7^2 - 0.00002485y_7^3 - 0.003644y_1y_8\\
&- 0.005698y_1^2y_8 + 0.1129y_2y_8 - 0.00029y_1y_2y_8 + 0.0119y_2^2y_8 - 0.02456y_3y_8 + 0.00513y_1y_3y_8 + 0.00451y_2y_3y_8\\
&- 0.0002579y_3^2y_8 - 0.03221y_4y_8 - 0.001087y_1y_4y_8 - 0.003224y_2y_4y_8 - 0.0005324y_3y_4y_8 - 0.0006357y_4^2y_8\\
&+ 0.0003821y_5y_8 + 0.001468y_1y_5y_8 - 0.003978y_2y_5y_8 - 0.0001739y_3y_5y_8 - 0.000335y_4y_5y_8 - 0.00009949y_5^2y_8\\
&+ 0.00007147y_6y_8 - 0.00001756y_1y_6y_8 - 0.0002214y_2y_6y_8 - 0.00001714y_3y_6y_8 - 0.00002434y_4y_6y_8\\
&- 9.68\times10^{-6}y_5y_6y_8 - 2.417\times10^{-7}y_6^2y_8 - 0.0144y_7y_8 + 0.003274y_1y_7y_8 - 0.0001942y_2y_7y_8 - 0.0004581y_3y_7y_8
\end{aligned}
$$

$$
\begin{aligned}
&- 0.0007273y_4y_7y_8 - 0.0001893y_5y_7y_8 - 0.00001319y_6y_7y_8 - 0.0002389y_7^2y_8 - 0.01242y_8^2 + 0.03921y_1y_8^2 \\
&+ 0.2146y_2y_8^2 + 0.02043y_3y_8^2 + 0.02525y_4y_8^2 + 0.001842y_5y_8^2 - 0.00001803y_6y_8^2 + 0.01295y_7y_8^2 + 0.03795y_8^3 \\
&- 0.168\alpha^2 - 0.001628y_1\alpha^2 - 0.005737y_2\alpha^2 - 0.00134y_3\alpha^2 + 0.002829y_4\alpha^2 - 0.0006488y_5\alpha^2 - 8.674{\times}10^{-6}y_6\alpha^2 \\
&+ 0.001117y_7\alpha^2 - 0.002095y_8\alpha^2
\end{aligned}
$$

$$
\begin{aligned}
x_4 = y_4 &- 0.01398y_1^2 - 0.001879y_1^3 - 0.04915y_1y_2 - 0.00494y_1^2y_2 - 0.02474y_2^2 - 0.003043y_1y_2^2 - 0.005346y_2^3 \\
&- 0.06293y_1y_3 - 0.03771y_1^2y_3 - 0.04952y_2y_3 + 0.001686y_1y_2y_3 - 0.001575y_3^2 + 0.0002738y_1y_3^2 + 0.00004003y_2y_3^2 \\
&+ 7.948{\times}10^{-6}y_3^3 + 0.09048y_1y_4 - 0.006031y_1^2y_4 - 0.02348y_2y_4 - 0.03631y_1y_2y_4 - 0.0009099y_3y_4 \\
&+ 0.0007665y_1y_3y_4 + 0.002155y_2y_3y_4 + 4.764{\times}10^{-6}y_3^2y_4 - 0.006183y_4^2 - 0.00159y_1y_4^2 + 0.0002275y_2y_4^2 \\
&+ 8.727{\times}10^{-7}y_3y_4^2 - 0.00003193y_4^3 - 0.02174y_1y_5 - 0.002656y_1^2y_5 + 0.001122y_2y_5 - 0.00268y_1y_2y_5 \\
&- 0.002064y_2^2y_5 - 0.001897y_3y_5 + 0.0003063y_1y_3y_5 + 0.0009095y_2y_3y_5 - 2.692{\times}10^{-6}y_3^2y_5 - 0.004765y_4y_5 \\
&- 0.0005045y_1y_4y_5 + 0.0001863y_2y_4y_5 + 0.00001588y_3y_4y_5 - 1.934{\times}10^{-7}y_4^2y_5 - 0.002589y_5^2 + 0.000368y_1y_5^2 \\
&- 0.0002181y_2y_5^2 + 4.381{\times}10^{-6}y_3y_5^2 - 0.00001074y_4y_5^2 - 7.45{\times}10^{-6}y_5^3 + 0.0006052y_1y_6 + 0.00004771y_1^2y_6 \\
&+ 0.0002462y_2y_6 + 0.00003016y_1y_2y_6 + 0.00001106y_2^2y_6 - 0.00002142y_3y_6 + 1.967{\times}10^{-6}y_1y_3y_6 \\
&+ 0.00001954y_2y_3y_6 + 1.031{\times}10^{-8}y_3^2y_6 - 0.00009455y_4y_6 - 0.00002015y_1y_4y_6 + 1.72{\times}10^{-6}y_2y_4y_6 \\
&+ 3.958{\times}10^{-7}y_3y_4y_6 + 1.235{\times}10^{-7}y_4^2y_6 - 0.00008833y_5y_6 + 0.00002592y_1y_5y_6 - 0.00001742y_2y_5y_6 \\
&+ 1.64{\times}10^{-8}y_3y_5y_6 - 1.532{\times}10^{-6}y_4y_5y_6 - 1.02{\times}10^{-6}y_5^2y_6 - 6.255{\times}10^{-7}y_6^2 + 8.588{\times}10^{-7}y_1y_6^2 - 5.559{\times}10^{-7}y_2y_6^2 \\
&- 5.585{\times}10^{-9}y_3y_6^2 - 5.575{\times}10^{-8}y_4y_6^2 - 4.477{\times}10^{-8}y_5y_6^2 - 5.033{\times}10^{-1}0y_6^3 + 0.06991y_1y_7 + 0.001902y_1^2y_7 \\
&- 0.04364y_2y_7 - 0.005323y_1y_2y_7 + 0.002109y_2^2y_7 - 0.0005085y_3y_7 + 0.0004046y_1y_3y_7 + 0.001078y_2y_3y_7 \\
&+ 0.0000118y_3^2y_7 - 0.003972y_4y_7 - 0.0006128y_1y_4y_7 + 0.0007717y_2y_4y_7 + 0.00001952y_3y_4y_7 - 0.00001992y_4^2y_7 \\
&- 0.002027y_5y_7 + 0.0001894y_1y_5y_7 + 0.0003707y_2y_5y_7 + 7.84{\times}10^{-6}y_3y_5y_7 - 6.699{\times}10^{-8}y_4y_5y_7 \\
&- 5.842{\times}10^{-6}y_5^2y_7 - 0.00003445y_6y_7 - 2.16{\times}10^{-7}y_1y_6y_7 + 5.328{\times}10^{-6}y_2y_6y_7 + 1.706{\times}10^{-7}y_3y_6y_7 \\
&- 1.419{\times}10^{-8}y_4y_6y_7 - 7.109{\times}10^{-7}y_5y_6y_7 - 2.316{\times}10^{-8}y_6^2y_7 - 0.0007788y_7^2 + 0.00001748y_1y_7^2 + 0.0002561y_2y_7^2 \\
&+ 9.248{\times}10^{-6}y_3y_7^2 - 6.036{\times}10^{-6}y_4y_7^2 - 6.594{\times}10^{-7}y_5y_7^2 - 3.222{\times}10^{-8}y_6y_7^2 + 2.789{\times}10^{-7}y_7^3 - 0.09614y_1y_8 \\
&- 0.006481y_1^2y_8 - 0.07578y_2y_8 - 0.01064y_1y_2y_8 - 0.01218y_2^2y_8 + 0.004531y_3y_8 + 0.0009064y_1y_3y_8 \\
&+ 0.004589y_2y_3y_8 + 0.00002726y_3^2y_8 - 0.02401y_4y_8 - 0.0004461y_1y_4y_8 + 0.002179y_2y_4y_8 + 0.0001066y_3y_4y_8 \\
&- 0.00006922y_4^2y_8 - 0.007327y_5y_8 + 0.003934y_1y_5y_8 - 0.0005284y_2y_5y_8 + 0.00005616y_3y_5y_8 - 0.00008422y_4y_5y_8 \\
&- 0.00008557y_5^2y_8 - 0.0001124y_6y_8 + 0.0001885y_1y_6y_8 - 0.00008975y_2y_6y_8 + 4.9{\times}10^{-8}y_3y_6y_8 \\
&- 8.671{\times}10^{-6}y_4y_6y_8 - 9.134{\times}10^{-6}y_5y_6y_8 - 1.968{\times}10^{-7}y_6^2y_8 - 0.0081y_7y_8 + 0.001241y_1y_7y_8 + 0.001322y_2y_7y_8 \\
&+ 0.00006402y_3y_7y_8 - 0.00008241y_4y_7y_8 - 0.00006365y_5y_7y_8 - 4.088{\times}10^{-6}y_6y_7y_8 - 0.00002006y_7^2y_8 \\
&- 0.02414y_8^2 - 0.182y_1y_8^2 + 0.1261y_2y_8^2 + 0.001514y_3y_8^2 + 0.01061y_4y_8^2 + 0.005221y_5y_8^2 + 0.00007374y_6y_8^2 \\
&+ 0.004079y_7y_8^2 + 0.02177y_8^3 + 0.1689\alpha^2 + 0.003014y_1\alpha^2 + 0.002441y_2\alpha^2 + 0.0004343y_5\alpha^2 - 4.701{\times}10^{-6}y_6\alpha^2 \\
&+ 0.002217y_7\alpha^2 + 0.00422y_8\alpha^2
\end{aligned}
$$

$$x_5 = y_5$$

$$x_6 = y_6$$

$$
\begin{aligned}
x_7 = y_7 &- 0.03866y_1^2 - 0.003413y_1^3 - 0.06335y_1y_2 - 0.006542y_1^2y_2 - 0.1317y_2^2 - 0.00828y_1y_2^2 - 0.01371y_2^3 - 0.1076y_1y_3 \\
&- 0.009591y_1^2y_3 - 0.3346y_2y_3 + 0.03926y_1y_2y_3 + 0.00458y_2^2y_3 + 0.01797y_3^2 - 0.002798y_1y_3^2 - 0.001253y_2y_3^2 \\
&+ 0.000287y_3^3 + 0.236y_1y_4 - 0.0319y_1^2y_4 + 0.003575y_2y_4 - 0.02201y_1y_2y_4 + 0.01759y_2^2y_4 + 0.02132y_3y_4 \\
&- 0.0001487y_1y_3y_4 - 0.007622y_2y_3y_4 + 0.000642y_3^2y_4 + 0.03172y_4^2 + 0.009387y_1y_4^2 + 0.001475y_2y_4^2 \\
&+ 0.000611y_3y_4^2 + 0.0003827y_4^3 - 0.01371y_1y_5 - 0.00366y_1^2y_5 - 0.01301y_2y_5 - 0.00311y_1y_2y_5 - 0.007541y_2^2y_5 \\
&+ 0.001113y_3y_5 - 0.0008713y_1y_3y_5 - 0.000551y_2y_3y_5 - 0.00003594y_3^2y_5 + 0.006997y_4y_5 - 0.0007145y_1y_4y_5 \\
&+ 0.0007672y_2y_4y_5 - 0.00004547y_3y_4y_5 + 0.00001581y_4^2y_5 + 0.007265y_5^2 - 0.0009977y_1y_5^2 + 0.0003618y_2y_5^2 \\
&+ 0.00006402y_3y_5^2 + 0.0001042y_4y_5^2 + 0.00001168y_5^3 + 0.00007289y_1y_6 + 0.00007279y_1^2y_6 - 0.0001411y_2y_6 \\
&+ 0.00003061y_1y_2y_6 + 0.0001349y_2^2y_6 - 0.00003337y_3y_6 + 0.00001517y_1y_3y_6 - 0.00005053y_2y_3y_6 \\
&- 1.574{\times}10^{-6}y_3^2y_6 + 0.000122y_4y_6 + 0.00005969y_1y_4y_6 + 3.284{\times}10^{-6}y_2y_4y_6 - 1.979{\times}10^{-6}y_3y_4y_6 \\
&- 1.703{\times}10^{-7}y_4^2y_6 + 0.0005265y_5y_6 - 0.0000846y_1y_5y_6 + 0.00001923y_2y_5y_6 + 6.698{\times}10^{-6}y_3y_5y_6 \\
&+ 0.00001144y_4y_5y_6 + 1.64{\times}10^{-6}y_5^2y_6 + 0.00001454y_6^2 - 2.6{\times}10^{-6}y_1y_6^2 + 2.112{\times}10^{-7}y_2y_6^2 + 2.071{\times}10^{-7}y_3y_6^2 \\
&+ 3.736{\times}10^{-7}y_4y_6^2 + 7.497{\times}10^{-8}y_5y_6^2 + 9.281{\times}10^{-1}0y_6^3 - 0.1564y_1y_7 + 0.008505y_1^2y_7 - 0.00386y_2y_7 \\
&+ 0.007362y_1y_2y_7 + 0.01324y_3y_7 - 0.003747y_1y_3y_7 - 0.005718y_2y_3y_7 + 0.0004172y_3^2y_7 + 0.02523y_4y_7 \\
&+ 0.004091y_1y_4y_7 - 0.002324y_2y_4y_7 + 0.0006758y_3y_4y_7 + 0.0005056y_4^2y_7 + 0.003422y_5y_7 - 0.001008y_1y_5y_7
\end{aligned}
$$

$$
\begin{aligned}
&+ 0.00006976y_2y_5y_7 - 0.00002346y_3y_5y_7 + 0.00002762y_4y_5y_7 + 0.00005865y_5^2y_7 + 0.00006843y_6y_7\\
&+ 7.523\times10^{-6}y_1y_6y_7 - 3.241\times10^{-6}y_2y_6y_7 - 7.901\times10^{-7}y_3y_6y_7 + 6.917\times10^{-7}y_4y_6y_7 + 6.091\times10^{-6}y_5y_6y_7\\
&+ 1.91\times10^{-7}y_6^2y_7 + 0.006318y_7^2 - 0.0006548y_1y_7^2 - 0.001309y_2y_7^2 + 0.0002191y_3y_7^2 + 0.0002659y_4y_7^2\\
&+ 0.00001029y_5y_7^2 + 3.709\times10^{-7}y_6y_7^2 + 0.00005266y_7^3 - 0.1212y_1y_8 - 0.008134y_1^2y_8 - 0.1528y_2y_8\\
&- 0.005201y_1y_2y_8 - 0.03059y_2^2y_8 - 0.0006423y_3y_8 - 0.0136y_1y_3y_8 - 0.01721y_2y_3y_8 + 0.0005924y_3^2y_8\\
&+ 0.09558y_4y_8 + 0.006536y_1y_4y_8 - 0.002785y_2y_4y_8 + 0.001061y_3y_4y_8 + 0.001373y_4^2y_8 + 0.03784y_5y_8\\
&- 0.009849y_1y_5y_8 - 0.0006358y_2y_5y_8 + 0.0002865y_3y_5y_8 + 0.0007586y_4y_5y_8 + 0.0002919y_5^2y_8 + 0.00245y_6y_8\\
&- 0.0003707y_1y_6y_8 - 9.541\times10^{-6}y_2y_6y_8 + 0.00002896y_3y_6y_8 + 0.00006004y_4y_6y_8 + 0.00003159y_5y_6y_8\\
&+ 8.827\times10^{-7}y_6^2y_8 + 0.05122y_7y_8 - 0.008303y_1y_7y_8 - 0.002967y_2y_7y_8 + 0.0008701y_3y_7y_8 + 0.001638y_4y_7y_8\\
&+ 0.0004565y_5y_7y_8 + 0.00003247y_6y_7y_8 + 0.0005347y_7^2y_8 - 2.601y_8^2 + 0.3988y_1y_8^2 + 0.005002y_2y_8^2 - 0.03491y_3y_8^2\\
&- 0.06572y_4y_8^2 - 0.007757y_5y_8^2 - 0.00006561y_6y_8^2 - 0.03226y_7y_8^2 - 0.137y_8^3 + 0.4128\alpha^2 + 0.003595y_1\alpha^2\\
&+ 0.01086y_2\alpha^2 + 0.007585y_3\alpha^2 - 0.005772y_4\alpha^2 + 0.001062y_5\alpha^2 - 6.543\times10^{-6}y_6\alpha^2 + 0.009658y_8\alpha^2
\end{aligned}
$$

$$
\begin{aligned}
x_8 = {}& y_8 + 0.01041y_1^2 + 0.0009975y_1^3 + 0.0125y_1y_2 + 0.001209y_1^2y_2 + 0.01822y_2^2 + 0.001596y_1y_2^2 + 0.001749y_2^3\\
&- 0.007779y_1y_3 - 0.001094y_1^2y_3 - 0.01174y_2y_3 - 0.00175y_1y_2y_3 + 0.00217y_2^2y_3 - 0.0009694y_3^2 + 0.0001936y_1y_3^2\\
&+ 0.0002007y_2y_3^2 - 0.00002157y_3^3 + 0.01451y_1y_4 + 0.001721y_1^2y_4 - 0.02039y_2y_4 - 0.002931y_1y_2y_4\\
&- 0.001181y_2^2y_4 - 0.001725y_3y_4 - 0.00019y_1y_3y_4 + 0.0003721y_2y_3y_4 - 0.0000442y_3^2y_4 - 0.001415y_4^2\\
&- 0.0005517y_1y_4^2 - 0.0001012y_2y_4^2 - 0.00004194y_3y_4^2 - 0.00002201y_4^3 + 0.002942y_1y_5 - 0.001744y_1^2y_5\\
&- 0.000896y_2y_5 + 0.000471y_1y_2y_5 - 0.001578y_2^2y_5 - 0.0002743y_3y_5 + 0.00008367y_1y_3y_5 + 0.00004695y_2y_3y_5\\
&+ 2.55\times10^{-6}y_3^2y_5 - 0.0004617y_4y_5 - 0.00001182y_1y_4y_5 - 0.00005445y_2y_4y_5 + 2.935\times10^{-6}y_3y_4y_5\\
&- 1.973\times10^{-6}y_4^2y_5 - 0.0002851y_5^2 + 0.00003543y_1y_5^2 - 0.00004446y_2y_5^2 - 4.7\times10^{-6}y_3y_5^2 - 6.292\times10^{-6}y_4y_5^2\\
&- 9.052\times10^{-7}y_5^3 + 0.00005457y_1y_6 - 0.00007761y_1^2y_6 + 0.0000285y_2y_6 - 3.839\times10^{-6}y_1y_2y_6 - 0.00007995y_2^2y_6\\
&- 2.901\times10^{-6}y_3y_6 - 7.949\times10^{-7}y_1y_3y_6 + 2.447\times10^{-6}y_2y_3y_6 + 1.132\times10^{-7}y_3^2y_6 - 7.802\times10^{-6}y_4y_6\\
&- 3.069\times10^{-6}y_1y_4y_6 + 4.114\times10^{-1}0y_2y_4y_6 + 1.242\times10^{-7}y_3y_4y_6 - 1.672\times10^{-8}y_4^2y_6 - 0.00001046y_5y_6\\
&+ 2.34\times10^{-6}y_1y_5y_6 - 3.609\times10^{-6}y_2y_5y_6 - 5.146\times10^{-7}y_3y_5y_6 - 6.581\times10^{-7}y_4y_5y_6 - 1.198\times10^{-7}y_5^2y_6\\
&- 1.158\times10^{-7}y_6^2 + 7.499\times10^{-8}y_1y_6^2 - 1.076\times10^{-7}y_2y_6^2 - 1.642\times10^{-8}y_3y_6^2 - 2.08\times10^{-8}y_4y_6^2 - 5.06\times10^{-9}y_5y_6^2\\
&+ 0.005481y_1y_7 - 0.0001009y_1^2y_7 - 0.006836y_2y_7 - 0.0001507y_1y_2y_7 + 0.0007276y_2^2y_7 - 0.001083y_3y_7\\
&+ 0.0002184y_1y_3y_7 + 0.0003326y_2y_3y_7 - 0.00003008y_3^2y_7 - 0.001379y_4y_7 - 0.0002434y_1y_4y_7 + 0.0001089y_2y_4y_7\\
&- 0.0000478y_3y_4y_7 - 0.00003177y_4^2y_7 - 0.0002323y_5y_7 + 0.0000542y_1y_5y_7 - 3.733\times10^{-6}y_2y_5y_7\\
&+ 1.641\times10^{-6}y_3y_5y_7 - 2.025\times10^{-6}y_4y_5y_7 - 3.689\times10^{-6}y_5^2y_7 - 3.601\times10^{-6}y_6y_7 - 5.856\times10^{-7}y_1y_6y_7\\
&+ 3.464\times10^{-7}y_2y_6y_7 + 4.978\times10^{-8}y_3y_6y_7 - 5.114\times10^{-8}y_4y_6y_7 - 3.781\times10^{-7}y_5y_6y_7 - 1.163\times10^{-8}y_6^2y_7\\
&- 0.0003785y_7^2 + 0.00004768y_1y_7^2 + 0.00008062y_2y_7^2 - 0.00001601y_3y_7^2 - 0.00001734y_4y_7^2 - 6.633\times10^{-7}y_5y_7^2\\
&- 2.408\times10^{-8}y_6y_7^2 - 3.645\times10^{-6}y_7^3 + 0.02835y_1y_8 - 0.00179y_1^2y_8 + 7.19\times10^{-6}y_2y_8 + 0.001652y_1y_2y_8\\
&- 0.002412y_3y_8 + 0.0007279y_1y_3y_8 + 0.0006777y_2y_3y_8 - 0.00003753y_3^2y_8 - 0.004693y_4y_8 - 0.0003127y_1y_4y_8\\
&- 0.0002523y_2y_4y_8 - 0.00007737y_3y_4y_8 - 0.00008105y_4^2y_8 - 0.0008526y_5y_8 + 0.0004095y_1y_5y_8\\
&- 0.0002994y_2y_5y_8 - 0.00002305y_3y_5y_8 - 0.0000446y_4y_5y_8 - 0.00001741y_5^2y_8 - 0.00001705y_6y_8\\
&+ 0.00001257y_1y_6y_8 - 0.0000169y_2y_6y_8 - 2.386\times10^{-6}y_3y_6y_8 - 3.371\times10^{-6}y_4y_6y_8 - 1.76\times10^{-6}y_5y_6y_8\\
&- 4.321\times10^{-8}y_6^2y_8 - 0.0023y_7y_8 + 0.0004566y_1y_7y_8 + 0.00004006y_2y_7y_8 - 0.00006551y_3y_7y_8 - 0.0001007y_4y_7y_8\\
&- 0.00002866y_5y_7y_8 - 1.97\times10^{-6}y_6y_7y_8 - 0.00003481y_7^2y_8 - 0.004691y_8^2 - 0.0137y_1y_8^2 + 0.01807y_2y_8^2\\
&+ 0.002855y_3y_8^2 + 0.003574y_4y_8^2 + 0.0005493y_5y_8^2 + 4.044\times10^{-6}y_6y_8^2 + 0.00193y_7y_8^2 + 0.006093y_8^3 - 0.07251\alpha^2\\
&- 0.001068y_1\alpha^2 - 0.00133y_2\alpha^2 + 7.653\times10^{-6}y_3\alpha^2 + 0.0008162y_4\alpha^2 + 0.0004083y_5\alpha^2\\
&+ 0.00002465y_6\alpha^2 + 0.0004339y_7\alpha^2. \qquad \text{(A.9)}
\end{aligned}
$$

Applying this change of variables to the Taylor approximation of Eq. (A.8) and substituting $\epsilon = \alpha^2$ the resulting 8-dimensional normal form system to third order is:

$$\dot{y_1} = 0.01384\epsilon y_1 + (65.20 + 0.08309\epsilon)y_2 - 0.01815y_1^3 - 0.1036y_1^2y_2 - 0.01815y_1y_2^2 - 0.1036y_2^3$$

$$\dot{y_2} = -(65.20 + 0.08309\epsilon)y_1 + 0.01384\epsilon y_2 + 0.1036y_1^3 - 0.01815y_1^2y_2 + 0.1036y_1y_2^2 - 0.01815y_2^3$$

$$
\begin{aligned}
\dot{y_3} &= -(129 + 0.1768\epsilon)y_3 + (78.75 - 0.1513\epsilon)y_4 - 0.1442y_1^2y_3 - 0.1442y_2^2y_3 + 0.08727y_1^2y_4 + 0.08727y_2^2y_4 \\
\dot{y_4} &= -(78.75 - 0.1513\epsilon)y_3 - (129 + 0.1768\epsilon)y_4 - 0.08727y_1^2y_3 - 0.08727y_2^2y_3 - 0.1442y_1^2y_4 - 0.1442y_2^2y_4 \\
\dot{y_5} &= -100y_5 + y_6 \\
\dot{y_6} &= -100y_6 \\
\dot{y_7} &= -(161.9 - 0.3975\epsilon)y_7 + 0.2186y_1^2y_7 + 0.2186y_2^2y_7 \\
\dot{y_8} &= -(80.14 + 0.07152\epsilon)y_8 + 0.1424y_1^2y_8 + 0.1424y_2^2y_8.
\end{aligned} \tag{A.10}
$$

Note that the equations for the bifurcating variables y_1 and y_2 are now decoupled from the other six equations. Expressing the (y_1, y_2) equations in polar coordinates,

$$
\begin{aligned}
\dot{r} &= 0.01384\epsilon r - 0.01815r^3 \\
\dot{\theta} &= -65.20 - 0.08309\epsilon + 0.1036r^2.
\end{aligned} \tag{A.11}
$$

In this form the circular symmetry (equivariance) of the normal form is clear, as the flow is θ-independent. Rescaling the radius by $R = \sqrt{0.01815}\, r$,

$$
\begin{aligned}
\dot{R} &= 0.01384\epsilon R - R^3 \\
\dot{\theta} &= -65.20 - 0.08309\epsilon + 5.706R^2,
\end{aligned} \tag{A.12}
$$

or in terms of a single complex variable $z = Re^{i\theta}$,

$$
\dot{z} = (-65.20i + (0.01384 - 0.08309i)\epsilon)z + (-1 + 5.706i)z|z|^2. \tag{A.13}
$$

A.3 Transformation of coupling

From the nonlinear coupling function $\mathbf{h}(\mathbf{v}_i, \mathbf{v}_j) : \mathbb{R}^8 \times \mathbb{R}^8 \to \mathbb{R}^8$ of the Jansen-Rit model, Eq. (3.42), we seek a first order approximation when both systems i and j are close to the bifurcation point H1,

expressed in terms of the normal form coordinate system. From Eq. (3.38) and Eq. (3.42),

$$\frac{\partial}{\partial \mathbf{v}_j}\mathbf{h}(\mathbf{v}_i,\mathbf{v}_j)=0\,,\qquad (\frac{\partial}{\partial \mathbf{v}_i}\mathbf{h}(\mathbf{v}_i,\mathbf{v}_j)).(\mathbf{v}_i-\mathbf{v}_c)=\begin{bmatrix}0\\0\\0\\0\\0\\H_e\kappa_e\gamma_1 S'(v_{c2}-v_{c3})((v_{i2}-v_{i3})-(v_{c2}-v_{c3}))\\0\\0\end{bmatrix}, \tag{A.14}$$

$$S'(v)=\frac{2e_0\rho_1\exp\left[\rho_1(\rho_2-v)\right]}{(1+\exp\left[\rho_1(\rho_2-v)\right])^2}\qquad\text{and}\qquad \mathbf{h}(\mathbf{v}_c,\mathbf{v}_c)=\begin{bmatrix}0\\0\\0\\0\\0\\H_e\kappa_e\gamma_1 S(v_{c2}-v_{c3})\\0\\0\end{bmatrix}, \tag{A.15}$$

which gives the following bilinear approximation of the coupling vector field near bifurcation point H1, expressed in the original Jansen-Rit coordinate system:

$$\begin{aligned}\tilde{\mathbf{h}}(\mathbf{v}_i,\mathbf{v}_j) &= \mathbf{h}(\mathbf{v}_c,\mathbf{v}_c)+(\frac{\partial}{\partial \mathbf{v}_i}\mathbf{h}(\mathbf{v}_i,\mathbf{v}_j)).(\mathbf{v}_i-\mathbf{v}_c)+(\frac{\partial}{\partial \mathbf{v}_j}\mathbf{h}(\mathbf{v}_i,\mathbf{v}_j)).(\mathbf{v}_j-\mathbf{v}_c)\\ &= \begin{bmatrix}0\\0\\0\\0\\0\\325(3.01+0.6708(-6.74+v_{i2}-v_{i3}))\gamma_1\\0\\0\end{bmatrix}.\end{aligned} \tag{A.16}$$

Now translating the point $\mathbf{v}_c$ to the origin and applying the linear transformation M of Eq. (A.6) by a change of coordinates $\mathbf{v} - \mathbf{v}_c = M\mathbf{x}$ as before, the coupling to first order in the 8-dimensional normal form coordinates is $\mathbf{k} : \mathbb{R}^8 \times \mathbb{R}^8 \to \mathbb{R}^8$ given by

$$\begin{aligned}\mathbf{k}(\mathbf{x}_i, \mathbf{x}_j) &= M^{-1}\tilde{\mathbf{h}}(M\mathbf{x}_i + \mathbf{v}_c, M\mathbf{x}_j + \mathbf{v}_c) \\ &= \begin{bmatrix} 4.463(3.01 + 0.1459x_{i1} - 0.06431x_{i2} + 0.0106x_{i3} + 0.03229x_{i4} - 0.008626x_{i7} - 0.001794x_{i8})\gamma_1 \\ -0.142(3.01 + 0.1459x_{i1} - 0.06431x_{i2} + 0.0106x_{i3} + 0.03229x_{i4} - 0.008626x_{i7} - 0.001794x_{i8})\gamma_1 \\ -34.97(3.01 + 0.1459x_{i1} - 0.06431x_{i2} + 0.0106x_{i3} + 0.03229x_{i4} - 0.008626x_{i7} - 0.001794x_{i8})\gamma_1 \\ 8.565(3.01 + 0.1459x_{i1} - 0.06431x_{i2} + 0.0106x_{i3} + 0.03229x_{i4} - 0.008626x_{i7} - 0.001794x_{i8})\gamma_1 \\ 0 \\ 0 \\ 66.82(3.01 + 0.1459x_{i1} - 0.06431x_{i2} + 0.0106x_{i3} + 0.03229x_{i4} - 0.008626x_{i7} - 0.001794x_{i8})\gamma_1 \\ -5.811(3.01 + 0.1459x_{i1} - 0.06431x_{i2} + 0.0106x_{i3} + 0.03229x_{i4} - 0.008626x_{i7} - 0.001794x_{i8})\gamma_1 \end{bmatrix}. \end{aligned} \tag{A.17}$$

As the eigenvalues corresponding to the normal form stable variables $x_{i3}, \cdots, x_{i8}$ in source region i all have real part < -80 near point H1, initial transients decay quickly leaving the stable variables in a range about zero that is small compared to center variables x_{i1} and x_{i2}. This justifies the reduction to two dimensions, expressing the first order effect of coupling on the oscillating center variables x_{j1} and x_{j2} as a function $\bar{\mathbf{k}} : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$,

$$\bar{\mathbf{k}}(\mathbf{x}_i, \mathbf{x}_j) = \gamma_1 \begin{bmatrix} 13.4 + 0.651x_{i1} - 0.287x_{i2} \\ -0.428 - 0.0207x_{i1} + 0.00913x_{i2} \end{bmatrix}, \tag{A.18}$$

or in terms of the complex variable $z_j = x_{j1} + \mathrm{i}x_{j2}$,

$$\begin{aligned} k(z_i, z_j) &= \gamma_1((13.4 - 0.428\mathrm{i}) + (0.651 - 0.0207\mathrm{i})\mathrm{Re}(z_i) + (-0.287 + 0.00913\mathrm{i})\mathrm{Im}(z_i)). \\ &= (1814 - 57.7\mathrm{i}) + (87.9 - 2.80\mathrm{i})\mathrm{Re}(z_i) + (-38.7 + 1.23\mathrm{i})\mathrm{Im}(z_i). \end{aligned} \tag{A.19}$$

Finally, if linearly rescaling the radius as in Eq. (A.13), $R = \sqrt{0.01815}r$, $z_i \to \sqrt{0.01815}z_i$, giving

$$k(z_i, z_j) = (244 - 7.77\mathrm{i}) + (87.9 - 2.80\mathrm{i})\mathrm{Re}(z_i) + (-38.7 + 1.23\mathrm{i})\mathrm{Im}(z_i). \tag{A.20}$$

To compute additional quadratic and cubic corrections to the coupling, one could go further and apply the near-identity transformation, Eq. (A.9) after linearly transforming a third order Taylor expansion of Eq. (3.42). But for this example we require only a leading order approximation to the coupling, and the linear terms here are nonzero.

Appendix B

Asymmetrical test system: construction and transformation

The non-trivial test system used in Section 5.3 was constructed by starting with the simple 3D example, Eq. (3.3), at the bifurcation point and breaking the symmetry of its nonlinear terms:

$$\begin{aligned} \dot{u}_1 &= -\omega u_2 - u_3 u_1 + \gamma u_1^2 + \beta_1 u_3 u_2 \\ \dot{u}_2 &= \omega u_1 - \alpha u_3 u_2 - \beta_2 u_3 u_1 \\ \dot{u}_3 &= -\lambda u_3 + (\kappa_1 u_1^2 + \kappa_2 u_2^2), \end{aligned} \tag{B.1}$$

where $\kappa_1 = 1.0, \kappa_2 = 4.0, \beta_1 = 2.0, \beta_2 = 3.0, \gamma = 0.7, \omega = 69.0, \lambda = 5.0$ and $\alpha = 0.5$, then applying a non-orthogonal linear transformation $S = R^{-1}QR$, where R is a rotation of $\frac{\pi}{8}$ about the vector $(1, 1, 0.5)^T$, i.e.,

$$R = \begin{bmatrix} \frac{4}{9} + \frac{5}{9}\cos\frac{\pi}{8} & \frac{4}{9} - \frac{4}{9}\cos\frac{\pi}{8} - \frac{1}{3}\sin\frac{\pi}{8} & \frac{2}{9} - \frac{2}{9}\cos\frac{\pi}{8} + \frac{2}{3}\sin\frac{\pi}{8} \\ \frac{4}{9} - \frac{4}{9}\cos\frac{\pi}{8} + \frac{1}{3}\sin\frac{\pi}{8} & \frac{4}{9} + \frac{5}{9}\cos\frac{\pi}{8} & \frac{2}{9} - \frac{2}{9}\cos\frac{\pi}{8} - \frac{2}{3}\sin\frac{\pi}{8} \\ \frac{2}{9} - \frac{2}{9}\cos\frac{\pi}{8} - \frac{2}{3}\sin\frac{\pi}{8} & \frac{2}{9} - \frac{2}{9}\cos\frac{\pi}{8} + \frac{2}{3}\sin\frac{\pi}{8} & \frac{1}{9} + \frac{8}{9}\cos\frac{\pi}{8} \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} 1 & 6 & 0.5 \\ 0 & 0 & 0.15 \\ 0 & 1.8 & 0 \end{bmatrix}, \tag{B.2}$$

before introducing an additive noise perturbation $(0, 0, \sigma dW)^T$ and deterministic perturbation $(0, -\epsilon, 0)^T$ to take the system through the bifurcation. Here is the resulting asymmetrical test system, to be used

as the starting point:

$$\begin{aligned}
dx_1 &= (311.479x_1 - 2201.82x_2 - 564.635x_3 - 20.0855x_1^2 \\
&\quad + 465.343x_2^2 + 65.9025x_1x_3 + 52.214x_1x_2 - 76.4356x_2x_3 - 54.2295x_3^2)\,dt \\
dx_2 &= (-2.21885x_1 + 30.7184x_2 + 6.63859x_3 + 0.255836x_1^2 \\
&\quad - 6.98162x_2^2 - 0.933232x_1x_3 - 1.40187x_1x_2 + 2.24106x_2x_3 + 0.87356x_3^2 - \epsilon)\,dt \\
dx_3 &= (195.124x_1 - 1312.67x_2 - 347.197x_3 - 12.4969x_1^2 \\
&\quad + 255.026x_2^2 + 40.7643x_1x_3 - 53.1318x_2x_3 + 33.9785x_1x_2 - 33.4977x_3^2)\,dt + \sigma\,dW.
\end{aligned} \tag{B.3}$$

Applying our method to this test system, we first set the noise intensity σ and bifurcation parameter ϵ to zero and compute the initial linear transformation:

$$\mathbf{x} = M\mathbf{w}, \qquad \text{where} \quad M = \begin{bmatrix} 3.2528 & 0 & -0.693649 \\ -0.0427981 & 0.0650435 & 0.105984 \\ 1.96129 & 0.143861 & -0.802078 \end{bmatrix}, \tag{B.4}$$

and near-identity normal form transformation at the bifurcation point:

$$\begin{aligned}
w_1 &= y_1 + 0.00406023y_1^2 - 0.000698326y_1^3 + 0.000117426y_1y_2 - 0.000154069y_1^2y_2 + 0.0017567y_2^2 \\
&\quad - 0.000832073y_1y_2^2 - 0.0000485379y_2^3 + 0.188591y_1y_3 + 0.00180971y_1^2y_3 - 0.646749y_2y_3 \\
&\quad - 0.00289789y_1y_2y_3 + 0.000666597y_2^2y_3 - 0.191593y_1y_3^2 - 0.122418y_2y_3^2 \\
w_2 &= y_2 + 0.000058713y_1^2 - 0.000066158y_1^3 + 0.00351341y_1y_2 + 0.000823212y_1^2y_2 + 0.00181501y_2^2 \\
&\quad - 0.000136449y_1y_2^2 + 0.000707187y_2^3 + 0.644715y_1y_3 + 0.00232999y_1^2y_3 + 0.198848y_2y_3 \\
&\quad + 0.00381244y_1y_2y_3 - 0.00139106y_2^2y_3 + 0.128518y_1y_3^2 - 0.187822y_2y_3^2 \\
w_3 &= y_3 + 0.138801y_1^2 - 0.0000315978y_1^3 + 0.00508242y_1y_2 + 0.000125656y_1^2y_2 + 0.142182y_2^2 \\
&\quad - 0.0000228631y_1y_2^2 + 0.000118876y_2^3 + 0.00252247y_1^2y_3 + 0.00642166y_1y_2y_3 \\
&\quad - 0.00252247y_2^2y_3.
\end{aligned} \tag{B.5}$$

Now applying these to transform the full system with noise and bifurcation parameter, after **step 1** (initial linear transformation $\mathbf{dw} = M^{-1}\underline{\mathbf{f}}(M\mathbf{w}, \epsilon)dt + \sigma M^{-1}G(M\mathbf{w}) \circ \mathbf{dW}$), we have:

$$\begin{aligned}
dw_1 &= (-69w_2 + 0.0121536w_1^2 - 0.075461w_1w_2 + 0.117133w_2^2 \\
&\quad - 1.08333w_1w_3 + 3.94144w_2w_3 - 0.788631\epsilon)\,dt - 0.356561\sigma \circ dW
\end{aligned}$$

$$
\begin{aligned}
dw_2 = (&69w_1 - 0.0377305w_1^2 + 0.234267w_1w_2 - 0.363638w_2^2 \\
&- 2.51588w_1w_3 - 0.853868w_2w_3 - 9.86728\epsilon)\, dt + 2.48988\sigma \circ dW \\
dw_3 = (&-5w_3 + 1.04469w_1^2 + 0.492006w_1w_2 + 0.360225w_2^2 - 3.69821\epsilon)\, dt \\
&- 1.67206\sigma \circ dW.
\end{aligned} \tag{B.6}
$$

Step 2 (applying the near identity transformation Eq. (B.5)) gives:

$$
\begin{aligned}
dy_1 = (&-69y_2 + \epsilon(-0.788631 + 0.705013y_1 - 2.35705y_2 - 6.23293y_3) \\
&- 0.135078y_1^3 + 0.454487y_1^2y_2 - 0.135078y_1y_2^2 + 0.454487y_2^3)\, dt \\
&+ (-0.356561 + 0.317939y_1 - 0.0163255y_1^2 - 1.09011y_2 + 0.193228y_1y_2 \\
&- 0.456581y_2^2 + 1.67757y_3 - 0.0141904y_1y_3 + 0.00258494y_2y_3 - 0.251419y_3^2)\sigma \circ dW \\
dy_2 = (&69y_1 + \epsilon(-9.86728 + 2.41905y_1 + 0.77397y_2 + 2.47053y_3) \\
&- 0.454487y_1^3 - 0.135078y_1^2y_2 - 0.454487y_1y_2^2 - 0.135078y_2^3)\, dt \\
&+ (2.48988 + 1.0693y_1 - 0.0576756y_1^2 + 0.3247y_2 + 0.439776y_1y_2 \\
&+ 0.135425y_2^2 - 0.265227y_3 + 0.00507589y_1y_3 + 0.0134997y_2y_3 - 0.515337y_3^2)\sigma \circ dW \\
dy_3 = (&-5y_3 + \epsilon(-3.69821 + 0.269076y_1 + 2.80992y_2) + 0.26815y_1^2y_3 + 0.26815y_2^2y_3)\, dt \\
&+ (-1.67206 + 0.0863277y_1 - 0.0898243y_1^2 - 0.706222y_2 + 0.00622183y_1y_2 \\
&- 0.0919066y_2^2 - 0.478541y_1y_3 + 0.0817461y_2y_3)\sigma \circ dW.
\end{aligned} \tag{B.7}
$$

Step 3 (approximate reduction to 2D and transformation to polar coordinates) gives:

$$
\begin{aligned}
dr = (&-0.135078r^3 - \epsilon(9.86728\sin\theta - 0.788631\cos\theta) \\
&+ \epsilon r(0.739492 + 0.0309986\sin 2\theta - 0.0344784\cos 2\theta))\, dt \\
&+ (2.48988\sin\theta - 0.356561\cos\theta + r(0.321319 - 0.0104069\sin 2\theta - 0.00338031\cos 2\theta) \\
&+ r^2(0.135457\sin\theta - 0.0164455\cos\theta + 0.0000318211\sin 3\theta + 0.000119974\cos 3\theta))\sigma \circ dW \\
d\theta = (&69 - 0.454487r^2 + \frac{\epsilon}{r}(0.788631\sin\theta - 9.86728\cos\theta) \\
&+ \epsilon(2.38805 + 0.0309986\cos 2\theta + 0.0344784\sin 2\theta))\, dt \\
&+ (\frac{1}{r}(0.356561\sin\theta + 2.48988\cos\theta) + 1.0797 - 0.0104069\cos 2\theta + 0.00338031\sin 2\theta \\
&+ r(0.456461\sin\theta - 0.0577074\cos\theta - 0.000119974\sin 3\theta + 0.0000318211\cos 3\theta))\sigma \circ dW.
\end{aligned} \tag{B.8}
$$

Step 4 (averaging the corresponding Fokker-Planck equation around the cycle and selecting a weakly equivalent SDE by Cholesky decomposition) gives:

$$
dr = (-0.135078r^3 + r(0.739492\epsilon + 0.0906116\sigma^2) + \frac{1}{r}1.58166\sigma^2)\, dt
$$

$$
\begin{aligned}
&\quad + (1.77857 + 0.125506r^2)\sigma \circ dW_A \\
d\theta &= (69 + 2.38805\epsilon - 0.17714\sigma^2 - 0.454487r^2)\, dt \\
&\quad + 0.522419r\sigma \circ dW_A + (\frac{1}{r}1.77857 + 0.333101r)\sigma \circ dW_B.
\end{aligned} \tag{B.9}
$$

Finally, linearly rescaling the radius by $R = \sqrt{0.135078}\, r$ puts the system in the standard form:

$$
\begin{aligned}
dR &= (-R^3 + R(0.739492\epsilon + 0.0906116\sigma^2) + \frac{1}{R}0.213647\sigma^2)\, dt \\
&\quad + (0.653677 + 0.341485R^2)\sigma \circ dW_A \\
d\theta &= (69 + 2.38805\epsilon - 0.17714\sigma^2 - 3.36464R^2)\, dt \\
&\quad + 1.42144R\sigma \circ dW_A + (\frac{1}{R}0.653677 + 0.906326R)\sigma \circ dW_B.
\end{aligned} \tag{B.10}
$$

N.B. If the extended system is used in step 2, deriving an ϵ-dependent transformation as discussed in Section 3.3.3, then the end result is the same as above but with third order corrections that show how the noise-induced effects vary through the bifurcation. For this system the corrections are insignificant:

$$
\begin{aligned}
dR &= (-R^3 + R(0.739492\epsilon + 0.0906116\sigma^2) + \frac{1}{R}(0.213647 + 0.000744846\epsilon)\sigma^2)\, dt \\
&\quad + (0.653677 + 0.00113947\epsilon + 0.341485R^2)\sigma \circ dW_A \\
d\theta &= (69 + 2.38805\epsilon - 0.17714\sigma^2 - 3.36464R^2)\, dt \\
&\quad + 1.42144R\sigma \circ dW_A + (\frac{1}{R}(0.653677 + 0.00113947\epsilon) + 0.906326R)\sigma \circ dW_B.
\end{aligned} \tag{B.11}
$$

Appendix C

Transforming the stochastic Jansen-Rit system

The Jansen-Rit model, with Gaussian white noise input to pyramidal and spiny stellate populations, was expressed as a system of first order Ito equations in Eqs (6.3)–(6.5).

C.1 Applying transformations to the SDEs

Step 1 (translation to the origin and linear transformation):
The bifurcation points are located and linear transformation M computed as shown in Appendix A.1, in the case of Hopf point H1 giving

$$p_c = 89.829\,\mathrm{s}^{-1}, \quad \mathbf{v}_c = (13.208, 20.165, 13.425, 3.3021, 0, 0, 0, 0)^{\mathrm{T}}\ \mathrm{mV}, \tag{C.1}$$

$$M = \begin{bmatrix} 0 & -0.49079 & -0.180728 & -0.110331 & 0 & 0 & -0.0988392 & -0.199643 \\ -0.0528038 & -0.0232788 & 0.0497681 & -0.00789791 & -0.246353 & -0.0123176 & -0.043254 & -0.848442 \\ -0.270252 & 0.0725844 & 0.0339665 & -0.0560267 & -0.246353 & -0.0123176 & -0.0303954 & -0.845768 \\ 0 & -0.122698 & -0.045182 & -0.0275829 & -0.04 & -0.0004 & -0.0247098 & -0.0499106 \\ 32 & 0 & 32 & 0 & 0 & 0 & 16 & 16 \\ 1.5178 & -3.44286 & -5.7976 & 4.93775 & 24.6353 & 0.985411 & 7.00193 & 67.9969 \\ -4.73257 & -17.6207 & 0.0305537 & 9.90153 & 24.6353 & 0.985411 & 4.92038 & 67.7825 \\ 8 & 0 & 8 & 0 & 4 & 0 & 4 & 4 \end{bmatrix}. \tag{C.2}$$

This time applying the translation and linear transformation $\mathbf{v} - \mathbf{v}_c = M\mathbf{x}$ to the full stochastic system Eq. (6.3), the linearly transformed system is then

$$\mathbf{dx} = M^{-1}\mathbf{f}(M\mathbf{x} + \mathbf{v}_c, \epsilon)dt + M^{-1}G(M\mathbf{x} + \mathbf{v}_c) \circ \mathbf{dW} \ , \tag{C.3}$$

which at Hopf point H1 gives the stochastic system,

$$\begin{aligned}
dx_1 &= (-1919.3 + 1/(0.0018319 + 0.0012107\exp(-0.12177x_1 + 0.053683x_2 - 0.0088489x_3 - 0.026952x_4 \\
&\quad + 0.0072008x_7 + 0.0014977x_8)) + 1/(-0.00023261 - 0.0010539\exp(0.068711x_2 + 0.025302x_3 + 0.015446x_4 \\
&\quad + 0.0224x_5 + 0.000224x_6 + 0.013837x_7 + 0.02795x_8)) + 1/(0.00041497 + 7.3273{\times}10^{-6}\exp(0.27484x_2 \\
&\quad + 0.10121x_3 + 0.061786x_4 + 0.05535x_7 + 0.1118x_8)) - 15.926x_1 + 39.762x_2 - 13.11x_3 - 10.822x_4 - 14.264x_5 \\
&\quad - 0.14264x_6 - 5.5954x_7 - 13.008x_8 + 4.4626\epsilon)\,dt + 4.4626\sigma_p \circ dW_2 \\
dx_2 &= (1475.9 + 1/(-0.00073893 - 0.00048835\exp(-0.12177x_1 + 0.053683x_2 - 0.0088489x_3 - 0.026952x_4 \\
&\quad + 0.0072008x_7 + 0.0014977x_8)) + 1/(-0.00030867 - 0.0013984\exp(0.068711x_2 + 0.025302x_3 + 0.015446x_4 \\
&\quad + 0.0224x_5 + 0.000224x_6 + 0.013837x_7 + 0.02795x_8)) + 1/(-0.013041 - 0.00023027\exp(0.27484x_2 \\
&\quad + 0.10121x_3 + 0.061786x_4 + 0.05535x_7 + 0.1118x_8)) - 25.72x_1 - 50.737x_2 - 9.4049x_3 + 1.2456x_4 - 10.749x_5 \\
&\quad - 0.10749x_6 - 9.0472x_7 - 14.044x_8 - 0.142\epsilon)\,dt - 0.142\sigma_p \circ dW_2 \\
dx_3 &= (14804. + 1/(0.00029395 + 0.00019427\exp(-0.12177x_1 + 0.053683x_2 - 0.0088489x_3 - 0.026952x_4 \\
&\quad + 0.0072008x_7 + 0.0014977x_8)) + 1/(0.0001062 + 0.00048115\exp(0.068711x_2 + 0.025302x_3 + 0.015446x_4 \\
&\quad + 0.0224x_5 + 0.000224x_6 + 0.013837x_7 + 0.02795x_8)) + 1/(-0.00005296 - 9.3514{\times}10^{-7}\exp(0.27484x_2 \\
&\quad + 0.10121x_3 + 0.061786x_4 + 0.05535x_7 + 0.1118x_8)) - 99.249x_1 + 51.105x_2 - 133.49x_3 + 58.431x_4 \\
&\quad + 31.242x_5 + 0.31242x_6 + 7.3493x_7 + 4.2109x_8 - 34.967\epsilon)\,dt - 34.967\sigma_p \circ dW_2 \\
dx_4 &= (-2943. + 1/(-0.00021213 - 0.0001402\exp(-0.12177x_1 + 0.053683x_2 - 0.0088489x_3 - 0.026952x_4 \\
&\quad + 0.0072008x_7 + 0.0014977x_8)) + 1/(0.00014621 + 0.0006624\exp(0.068711x_2 + 0.025302x_3 + 0.015446x_4 \\
&\quad + 0.0224x_5 + 0.000224x_6 + 0.013837x_7 + 0.02795x_8)) + 1/(0.00021622 + 3.8179{\times}10^{-6}\exp(0.27484x_2 \\
&\quad + 0.10121x_3 + 0.061786x_4 + 0.05535x_7 + 0.1118x_8)) + 137.53x_1 + 30.652x_2 - 35.138x_3 - 78.029x_4 \\
&\quad + 22.693x_5 + 0.22693x_6 + 10.25x_7 + 35.44x_8 + 8.5645\epsilon)\,dt + 8.5645\sigma_p \circ dW_2 \\
dx_5 &= (-100x_5 + x_6)\,dt \\
dx_6 &= -100x_6\,dt \\
dx_7 &= (-36878. + 1/(0.00017042 + 0.00011263\exp(-0.12177x_1 + 0.053683x_2 - 0.0088489x_3 - 0.026952x_4 \\
&\quad + 0.0072008x_7 + 0.0014977x_8)) + 1/(-0.000085655 - 0.00038807\exp(0.068711x_2 + 0.025302x_3 \\
&\quad + 0.015446x_4 + 0.0224x_5 + 0.000224x_6 + 0.013837x_7 + 0.02795x_8)) + 1/(0.000027714 \\
&\quad + 4.8936{\times}10^{-7}\exp(0.27484x_2 + 0.10121x_3 + 0.061786x_4 + 0.05535x_7 + 0.1118x_8)) - 171.18x_1 + 125.73x_2 \\
&\quad + 6.07x_3 - 26.589x_4 - 38.735x_5 - 0.38735x_6 - 141.63x_7 + 22.552x_8 + 66.819\epsilon)\,dt + 66.819\sigma_p \circ dW_2 \\
dx_8 &= (2854.8 + 1/(-0.01903 - 0.012577\exp(-0.12177x_1 + 0.053683x_2 - 0.0088489x_3 - 0.026952x_4 + 0.0072008x_7 \\
&\quad + 0.0014977x_8)) + 1/(0.0006943 + 0.0031456\exp(0.068711x_2 + 0.025302x_3 + 0.015446x_4 + 0.0224x_5 \\
&\quad + 0.000224x_6 + 0.013837x_7 + 0.02795x_8)) + 1/(-0.00031867 - 5.6269{\times}10^{-6}\exp(0.27484x_2 + 0.10121x_3 \\
&\quad + 0.061786x_4 + 0.05535x_7 + 0.1118x_8)) + 1.533x_1 - 0.7222x_2 + 0.094334x_3 + 0.32889x_4 + 4.7787x_5 \\
&\quad + 0.047787x_6 - 0.09999x_7 - 80.181x_8 - 5.8111\epsilon)\,dt - 5.8111\sigma_p \circ dW_2.
\end{aligned} \tag{C.4}$$

Step 2 (near-identity normal form transformation)

Both the deterministic terms and noise coefficient functions of Eq. (C.4) are now approximated near the origin in phase space and parameter space by Taylor series truncated to third order. In this and all subsequent series approximations the asymptotic scaling $O(\sigma_i^2) = O(x_i^2) = O(\epsilon)$ is used when

truncating as discussed in Section 4.4.1.

Next setting noise intensity σ_p to zero, the required near-identity transformation $\mathbf{x} = \mathbf{y} + \mathbf{q}(\mathbf{y})$ is computed to third order, similarly to Appendix A.2. To save space here we will display the intermediate steps from the ϵ-independent transformation, computed from the deterministic system exactly at bifurcation. The results from the full ϵ-dependent transformation were computed similarly. Note that the two methods give the same answer for the deterministic terms, as the averaging done in step 4 consolidates non-symmetrical ϵ-dependent terms to arrive at the same value of γ_ϵ. The latter method allows to see how noise effects vary through the bifurcation. Exactly at bifurcation, the near-identity transformation to third order is:

$$
\begin{aligned}
x_1 = {} & y_1 - 0.002530y_1^2 - 0.0007372y_1^3 - 0.02128y_1y_2 + 0.000455y_1^2y_2 + 0.00116y_2^2 - 0.001604y_1y_2^2 \\
& - 0.00003821y_2^3 - 0.00199y_1y_3 + 0.0002932y_1^2y_3 + 0.006288y_2y_3 - 0.000484y_1y_2y_3 - 0.0013y_2^2y_3 \\
& + 0.0007812y_3^2 - 0.0001419y_1y_3^2 - 0.0002229y_2y_3^2 + 0.00001517y_3^3 - 0.001713y_1y_4 - 0.00006445y_1^2y_4 \\
& + 0.008011y_2y_4 + 0.001762y_1y_2y_4 + 0.0001973y_2^2y_4 + 0.001403y_3y_4 + 0.0002499y_1y_3y_4 - 0.0002627y_2y_3y_4 \\
& + 0.00002713y_3^2y_4 + 0.0008333y_4^2 + 0.000395y_1y_4^2 + 0.00003935y_2y_4^2 + 0.000026y_3y_4^2 + 0.00001235y_4^3 \\
& - 0.00002612y_1y_5 + 0.0001229y_1^2y_5 + 0.003148y_2y_5 - 0.0003958y_1y_2y_5 + 0.0002887y_2^2y_5 + 0.0006217y_3y_5 \\
& - 0.0000877y_1y_3y_5 - 0.0001041y_2y_3y_5 - 1.762{\times}10^{-6}y_3^2y_5 + 0.0005858y_4y_5 + 0.0001152y_1y_4y_5 + 0.00002166y_2y_4y_5 \\
& - 3.496{\times}10^{-6}y_3y_4y_5 + 2.213{\times}10^{-6}y_4^2y_5 + 0.0003477y_5^2 - 0.00002571y_1y_5^2 + 0.000044y_2y_5^2 + 2.85{\times}10^{-6}y_3y_5^2 \\
& + 3.987{\times}10^{-6}y_4y_5^2 + 1.581{\times}10^{-6}y_5^3 + 0.00001439y_1y_6 - 3.104{\times}10^{-7}y_1^2y_6 + 0.00005133y_2y_6 + 3.091{\times}10^{-6}y_1y_2y_6 \\
& + 6.329{\times}10^{-7}y_2^2y_6 + 7.861{\times}10^{-6}y_3y_6 + 4.628{\times}10^{-7}y_1y_3y_6 - 2.408{\times}10^{-6}y_2y_3y_6 - 9.131{\times}10^{-8}y_3^2y_6 + 7.621{\times}10^{-6}y_4y_6 \\
& + 3.038{\times}10^{-6}y_1y_4y_6 - 2.758{\times}10^{-7}y_2y_4y_6 - 1.007{\times}10^{-7}y_3y_4y_6 + 2.315{\times}10^{-8}y_4^2y_6 + 8.63{\times}10^{-6}y_5y_6 \\
& - 5.801{\times}10^{-7}y_1y_5y_6 + 2.481{\times}10^{-6}y_2y_5y_6 + 3.945{\times}10^{-7}y_3y_5y_6 + 4.257{\times}10^{-7}y_4y_5y_6 + 1.99{\times}10^{-7}y_5^2y_6 \\
& + 5.097{\times}10^{-8}y_6^2 - 1.675{\times}10^{-8}y_1y_6^2 + 6.849{\times}10^{-8}y_2y_6^2 + 1.359{\times}10^{-8}y_3y_6^2 + 1.332{\times}10^{-8}y_4y_6^2 + 7.871{\times}10^{-9}y_5y_6^2 \\
& - 0.001413y_1y_7 + 0.0002356y_1^2y_7 + 0.004485y_2y_7 - 0.0001436y_1y_2y_7 - 0.0005201y_2^2y_7 + 0.0009191y_3y_7 \\
& - 0.0001503y_1y_3y_7 - 0.0002847y_2y_3y_7 + 0.00001962y_3^2y_7 + 0.0008702y_4y_7 + 0.0001841y_1y_4y_7 - 0.0001075y_2y_4y_7 \\
& + 0.00002927y_3y_4y_7 + 0.0000181y_4^2y_7 + 0.000368y_5y_7 - 0.00005045y_1y_5y_7 - 0.00004089y_2y_5y_7 - 2.213{\times}10^{-6}y_3y_5y_7 \\
& + 1.564{\times}10^{-6}y_4y_5y_7 + 2.829{\times}10^{-6}y_5^2y_7 + 4.597{\times}10^{-6}y_6y_7 + 2.319{\times}10^{-7}y_1y_6y_7 - 8.693{\times}10^{-7}y_2y_6y_7 \\
& - 4.806{\times}10^{-8}y_3y_6y_7 + 4.053{\times}10^{-8}y_4y_6y_7 + 2.938{\times}10^{-7}y_5y_6y_7 + 8.813{\times}10^{-9}y_6^2y_7 + 0.0002826y_7^2 \\
& - 0.00004164y_1y_7^2 - 0.00007229y_2y_7^2 + 0.0000102y_3y_7^2 + 0.00001081y_4y_7^2 + 5.042{\times}10^{-7}y_5y_7^2 + 1.926{\times}10^{-8}y_6y_7^2 \\
& + 2.358{\times}10^{-6}y_7^3 - 0.003178y_1y_8 + 0.0009023y_1^2y_8 + 0.01433y_2y_8 - 0.002151y_1y_2y_8 + 0.0006395y_2^2y_8 \\
& + 0.002276y_3y_8 - 0.0004707y_1y_3y_8 - 0.0004133y_2y_3y_8 + 0.00001552y_3^2y_8 + 0.002454y_4y_8 + 0.0002549y_1y_4y_8 \\
& + 0.0001709y_2y_4y_8 + 0.00004251y_3y_4y_8 + 0.00004155y_4^2y_8 + 0.0009084y_5y_8 - 0.0002351y_1y_5y_8 + 0.0002266y_2y_5y_8 \\
& + 0.00001147y_3y_5y_8 + 0.00002897y_4y_5y_8 + 0.00001851y_5^2y_8 + 0.00001107y_6y_8 - 3.537{\times}10^{-6}y_1y_6y_8 \\
& + 0.00001116y_2y_6y_8 + 1.952{\times}10^{-6}y_3y_6y_8 + 2.165{\times}10^{-6}y_4y_6y_8 + 1.768{\times}10^{-6}y_5y_6y_8 + 3.343{\times}10^{-8}y_6^2y_8 \\
& + 0.001438y_7y_8 - 0.000302y_1y_7y_8 - 0.00004886y_2y_7y_8 + 0.00003913y_3y_7y_8 + 0.00005972y_4y_7y_8 \\
& + 0.00002395y_5y_7y_8 + 1.504{\times}10^{-6}y_6y_7y_8 + 0.00002302y_7^2y_8 + 0.001934y_8^2 + 0.003221y_1y_8^2 - 0.01174y_2y_8^2 \\
& - 0.002437y_3y_8^2 - 0.002263y_4y_8^2 - 0.0009338y_5y_8^2 - 8.683{\times}10^{-6}y_6y_8^2 - 0.00145y_7y_8^2 - 0.003802y_8^3
\end{aligned}
$$

$$
\begin{aligned}
x_2 = {} & y_2 + 0.02432y_1^2 - 0.0005537y_1^3 - 0.01006y_1y_2 + 0.0004057y_1^2y_2 + 0.017y_2^2 + 0.00016y_1y_3 + 0.0005278y_1^2y_3 \\
& + 0.00346y_2y_3 + 0.0002419y_1y_2y_3 - 0.001208y_2^2y_3 + 0.0003411y_3^2 - 0.00006238y_1y_3^2 - 0.000145y_2y_3^2 \\
& + 2.555{\times}10^{-6}y_3^3 - 0.005681y_1y_4 - 0.0007362y_1^2y_4 + 0.006387y_2y_4 + 0.001571y_1y_2y_4 - 0.0001034y_2^2y_4 \\
& + 0.0006598y_3y_4 + 0.00009545y_1y_3y_4 - 0.0002642y_2y_3y_4 + 3.504{\times}10^{-6}y_3^2y_4 + 0.0004511y_4^2 + 0.0002183y_1y_4^2 \\
& - 0.00002776y_2y_4^2 + 3.666{\times}10^{-6}y_3y_4^2 + 4.469{\times}10^{-6}y_4^3 - 0.003108y_1y_5 + 0.0003779y_1^2y_5 + 0.004741y_2y_5 \\
& - 0.000288y_1y_2y_5 + 0.0002507y_2^2y_5 + 0.0006427y_3y_5 - 0.00009049y_1y_3y_5 - 0.0001826y_2y_3y_5 + 3.333{\times}10^{-8}y_3^2y_5 \\
& + 0.0008507y_4y_5 + 0.0001916y_1y_4y_5 - 0.00003535y_2y_4y_5 - 3.285{\times}10^{-6}y_3y_4y_5 + 2.583{\times}10^{-6}y_4^2y_5 + 0.0004976y_5^2 \\
& - 0.00003073y_1y_5^2 + 0.00003662y_2y_5^2 + 1.151{\times}10^{-8}y_3y_5^2 + 2.566{\times}10^{-6}y_4y_5^2 + 1.836{\times}10^{-6}y_5^3 - 0.0000738y_1y_6 \\
& + 2.282{\times}10^{-6}y_1^2y_6 + 0.00007991y_2y_6 + 4.35{\times}10^{-7}y_1y_2y_6 + 3.149{\times}10^{-6}y_2^2y_6 + 8.058{\times}10^{-6}y_3y_6 - 6.436{\times}10^{-7}y_1y_3y_6
\end{aligned}
$$

$$
\begin{aligned}
&- 3.27\times10^{-6}y_2y_3y_6 - 2.781\times10^{-8}y_3^2y_6 + 0.00001328y_4y_6 + 3.724\times10^{-6}y_1y_4y_6 - 8.802\times10^{-7}y_2y_4y_6\\
&- 7.706\times10^{-8}y_3y_4y_6 + 3.568\times10^{-8}y_4^2y_6 + 0.00001547y_5y_6 - 5.159\times10^{-7}y_1y_5y_6 + 1.712\times10^{-6}y_2y_5y_6\\
&+ 1.237\times10^{-7}y_3y_5y_6 + 2.686\times10^{-7}y_4y_5y_6 + 2.387\times10^{-7}y_5^2y_6 + 1.324\times10^{-7}y_6^2 - 1.316\times10^{-8}y_1y_6^2 + 4.401\times10^{-8}y_2y_6^2\\
&+ 5.59\times10^{-9}y_3y_6^2 + 8.41\times10^{-9}y_4y_6^2 + 9.847\times10^{-9}y_5y_6^2 - 0.001335y_1y_7 - 0.00005327y_1^2y_7 + 0.002791y_2y_7\\
&+ 0.0002029y_1y_2y_7 - 0.000487y_2^2y_7 + 0.000413y_3y_7 - 0.0000712y_1y_3y_7 - 0.0001964y_2y_3y_7 + 1.612\times10^{-6}y_3^2y_7\\
&+ 0.0005613y_4y_7 + 0.0001156y_1y_4y_7 - 0.0001168y_2y_4y_7 + 2.487\times10^{-6}y_3y_4y_7 + 5.902\times10^{-6}y_4^2y_7 + 0.0004542y_5y_7\\
&- 0.00003057y_1y_5y_7 - 0.00008036y_2y_5y_7 - 1.975\times10^{-6}y_3y_5y_7 + 5.92\times10^{-7}y_4y_5y_7 + 1.08\times10^{-6}y_5^2y_7\\
&+ 6.505\times10^{-6}y_6y_7 + 2.102\times10^{-7}y_1y_6y_7 - 1.272\times10^{-6}y_2y_6y_7 - 3.506\times10^{-8}y_3y_6y_7 + 1.702\times10^{-8}y_4y_6y_7\\
&+ 1.366\times10^{-7}y_5y_6y_7 + 4.244\times10^{-9}y_6^2y_7 + 0.0001362y_7^2 - 8.513\times10^{-6}y_1y_7^2 - 0.00005474y_2y_7^2 + 4.208\times10^{-7}y_3y_7^2\\
&+ 1.369\times10^{-6}y_4y_7^2 + 6.518\times10^{-8}y_5y_7^2 + 6.046\times10^{-9}y_6y_7^2 + 2.375\times10^{-7}y_7^3 - 0.0105y_1y_8 + 0.001537y_1^2y_8\\
&+ 0.009908y_2y_8 - 0.0006049y_1y_2y_8 + 0.0001646y_2^2y_8 + 0.001021y_3y_8 - 0.000267y_1y_3y_8 - 0.0004359y_2y_3y_8\\
&- 8.214\times10^{-6}y_3^2y_8 + 0.001753y_4y_8 + 0.0002523y_1y_4y_8 - 0.0001227y_2y_4y_8 - 0.00001105y_3y_4y_8 + 0.00001589y_4^2y_8\\
&+ 0.001393y_5y_8 - 0.0002351y_1y_5y_8 + 0.00007941y_2y_5y_8 - 5.562\times10^{-6}y_3y_5y_8 + 0.0000162y_4y_5y_8 + 0.00001726y_5^2y_8\\
&+ 0.00002239y_6y_8 - 4.504\times10^{-6}y_1y_6y_8 + 6.078\times10^{-6}y_2y_6y_8 + 7.189\times10^{-7}y_3y_6y_8 + 1.336\times10^{-6}y_4y_6y_8\\
&+ 1.696\times10^{-6}y_5y_6y_8 + 2.843\times10^{-8}y_6^2y_8 + 0.0008417y_7y_8 - 0.0001216y_1y_7y_8 - 0.0001374y_2y_7y_8\\
&- 6.964\times10^{-7}y_3y_7y_8 + 0.00001382y_4y_7y_8 + 0.000013y_5y_7y_8 + 7.451\times10^{-7}y_6y_7y_8 + 4.458\times10^{-6}y_7^2y_8 + 0.001483y_8^2\\
&+ 0.003305y_1y_8^2 - 0.007539y_2y_8^2 - 0.00112y_3y_8^2 - 0.001487y_4y_8^2 - 0.001189y_5y_8^2 - 0.00001537y_6y_8^2\\
&- 0.0007148y_7y_8^2 - 0.002247y_8^3
\end{aligned}
$$

$$
\begin{aligned}
x_3 = y_3 &+ 0.003911y_1^2 + 0.0002712y_1^3 + 0.04354y_1y_2 + 0.003925y_1^2y_2 + 0.06832y_2^2 + 0.001856y_1y_2^2 + 0.007315y_2^3\\
&+ 0.05264y_1y_3 - 0.003647y_1^2y_3 + 0.01463y_2y_3 + 0.02729y_1y_2y_3 - 0.001681y_2^2y_3 - 0.005696y_3^2 + 0.001454y_1y_3^2\\
&+ 0.001654y_2y_3^2 - 0.0001454y_3^3 + 0.0531y_1y_4 + 0.01091y_1^2y_4 - 0.03733y_2y_4 + 0.003773y_1y_2y_4 + 0.03061y_2^2y_4\\
&- 0.01284y_3y_4 - 0.001537y_1y_3y_4 + 0.002616y_2y_3y_4 - 0.0003067y_3^2y_4 - 0.009019y_4^2 - 0.00362y_1y_4^2\\
&- 0.0009328y_2y_4^2 - 0.000297y_3y_4^2 - 0.0001644y_4^3 + 0.008933y_1y_5 - 0.001089y_1^2y_5 + 0.03023y_2y_5 + 0.002246y_1y_2y_5\\
&+ 0.002592y_2^2y_5 - 0.002519y_3y_5 + 0.0004502y_1y_3y_5 - 0.00003444y_2y_3y_5 + 0.0000184y_3^2y_5 - 0.001312y_4y_5\\
&+ 0.0002854y_1y_4y_5 - 0.0005966y_2y_4y_5 + 0.0000207y_3y_4y_5 - 0.00001757y_4^2y_5 - 0.0002744y_5^2 + 0.00003333y_1y_5^2\\
&- 0.0003826y_2y_5^2 - 0.00003373y_3y_5^2 - 0.00004816y_4y_5^2 - 4.178\times10^{-6}y_5^3 + 0.0002572y_1y_6 + 0.000052y_1^2y_6\\
&- 0.0004796y_2y_6 - 0.00004784y_1y_2y_6 - 2.237\times10^{-7}y_2^2y_6 - 0.00005112y_3y_6 - 9.739\times10^{-6}y_1y_3y_6\\
&+ 5.104\times10^{-6}y_2y_3y_6 + 7.876\times10^{-7}y_3^2y_6 - 9.728\times10^{-6}y_4y_6 - 9.437\times10^{-6}y_1y_4y_6 - 4.888\times10^{-6}y_2y_4y_6\\
&+ 9.15\times10^{-7}y_3y_4y_6 - 2.113\times10^{-7}y_4^2y_6 + 0.00002731y_5y_6 - 7.609\times10^{-7}y_1y_5y_6 - 0.00003398y_2y_5y_6\\
&- 3.658\times10^{-6}y_3y_5y_6 - 4.832\times10^{-6}y_4y_5y_6 - 5.067\times10^{-7}y_5^2y_6 + 7.637\times10^{-7}y_6^2 - 3.177\times10^{-8}y_1y_6^2 - 1.091\times10^{-6}y_2y_6^2\\
&- 1.168\times10^{-7}y_3y_6^2 - 1.493\times10^{-7}y_4y_6^2 - 1.868\times10^{-8}y_5y_6^2 - 1.285\times10^{-1}0y_6^3 - 0.009502y_1y_7 - 0.00304y_1^2y_7\\
&- 0.08187y_2y_7 - 0.001779y_1y_2y_7 + 0.0048y_2^2y_7 - 0.007743y_3y_7 + 0.001486y_1y_3y_7 + 0.002232y_2y_3y_7\\
&- 0.0002026y_3^2y_7 - 0.009775y_4y_7 - 0.001547y_1y_4y_7 + 0.0005854y_2y_4y_7 - 0.0003324y_3y_4y_7 - 0.0002372y_4^2y_7\\
&- 0.0008755y_5y_7 + 0.0004192y_1y_5y_7 - 0.0001769y_2y_5y_7 + 0.0000105y_3y_5y_7 - 0.0000149y_4y_5y_7 - 0.00002455y_5^2y_7\\
&- 6.94\times10^{-6}y_6y_7 - 2.969\times10^{-6}y_1y_6y_7 - 7.946\times10^{-7}y_2y_6y_7 + 3.274\times10^{-7}y_3y_6y_7 - 3.753\times10^{-7}y_4y_6y_7\\
&- 2.533\times10^{-6}y_5y_6y_7 - 7.82\times10^{-8}y_6^2y_7 - 0.002545y_7^2 + 0.0003205y_1y_7^2 + 0.0005389y_2y_7^2 - 0.0001087y_3y_7^2\\
&- 0.0001206y_4y_7^2 - 4.207\times10^{-6}y_5y_7^2 - 1.57\times10^{-7}y_6y_7^2 - 0.00002485y_7^3 - 0.003644y_1y_8 - 0.005698y_1^2y_8\\
&+ 0.1129y_2y_8 - 0.00029y_1y_2y_8 + 0.0119y_2^2y_8 - 0.02456y_3y_8 + 0.00513y_1y_3y_8 + 0.00451y_2y_3y_8 - 0.0002579y_3^2y_8\\
&- 0.03221y_4y_8 - 0.001087y_1y_4y_8 - 0.003224y_2y_4y_8 - 0.0005324y_3y_4y_8 - 0.0006357y_4^2y_8 + 0.0003821y_5y_8\\
&+ 0.001468y_1y_5y_8 - 0.003978y_2y_5y_8 - 0.0001739y_3y_5y_8 - 0.000335y_4y_5y_8 - 0.00009949y_5^2y_8 + 0.00007147y_6y_8\\
&- 0.00001756y_1y_6y_8 - 0.0002214y_2y_6y_8 - 0.00001714y_3y_6y_8 - 0.00002434y_4y_6y_8 - 9.68\times10^{-6}y_5y_6y_8\\
&- 2.417\times10^{-7}y_6^2y_8 - 0.0144y_7y_8 + 0.003274y_1y_7y_8 - 0.0001942y_2y_7y_8 - 0.0004581y_3y_7y_8 - 0.0007273y_4y_7y_8\\
&- 0.0001893y_5y_7y_8 - 0.00001319y_6y_7y_8 - 0.0002389y_7^2y_8 - 0.01242y_8^2 + 0.03921y_1y_8^2 + 0.2146y_2y_8^2 + 0.02043y_3y_8^2\\
&+ 0.02525y_4y_8^2 + 0.001842y_5y_8^2 - 0.00001803y_6y_8^2 + 0.01295y_7y_8^2 + 0.03795y_8^3
\end{aligned}
$$

$$
\begin{aligned}
x_4 = y_4 &- 0.01398y_1^2 - 0.001879y_1^3 - 0.04915y_1y_2 - 0.00494y_1^2y_2 - 0.02474y_2^2 - 0.003043y_1y_2^2 - 0.005346y_2^3\\
&- 0.06293y_1y_3 - 0.03771y_1^2y_3 - 0.04952y_2y_3 + 0.001686y_1y_2y_3 - 0.001575y_3^2 + 0.0002738y_1y_3^2 + 0.00004003y_2y_3^2\\
&+ 7.948\times10^{-6}y_3^3 + 0.09048y_1y_4 - 0.006031y_1^2y_4 - 0.02348y_2y_4 - 0.03631y_1y_2y_4 - 0.0009099y_3y_4\\
&+ 0.0007665y_1y_3y_4 + 0.002155y_2y_3y_4 + 4.764\times10^{-6}y_3^2y_4 - 0.006183y_4^2 - 0.00159y_1y_4^2 + 0.0002275y_2y_4^2\\
&+ 8.727\times10^{-7}y_3y_4^2 - 0.00003193y_4^3 - 0.02174y_1y_5 - 0.002656y_1^2y_5 + 0.001122y_2y_5 - 0.00268y_1y_2y_5\\
&- 0.002064y_2^2y_5 - 0.001897y_3y_5 + 0.0003063y_1y_3y_5 + 0.0009095y_2y_3y_5 - 2.692\times10^{-6}y_3^2y_5 - 0.004765y_4y_5
\end{aligned}
$$

$$
\begin{aligned}
&- 0.0005045y_1y_4y_5 + 0.0001863y_2y_4y_5 + 0.00001588y_3y_4y_5 - 1.934\times10^{-7}y_4^2y_5 - 0.002589y_5^2 + 0.000368y_1y_5^2 \\
&- 0.0002181y_2y_5^2 + 4.381\times10^{-6}y_3y_5^2 - 0.00001074y_4y_5^2 - 7.45\times10^{-6}y_5^3 + 0.0006052y_1y_6 + 0.00004771y_1^2y_6 \\
&+ 0.0002462y_2y_6 + 0.00003016y_1y_2y_6 + 0.00001106y_2^2y_6 - 0.00002142y_3y_6 + 1.967\times10^{-6}y_1y_3y_6 \\
&+ 0.00001954y_2y_3y_6 + 1.031\times10^{-8}y_3^2y_6 - 0.00009455y_4y_6 - 0.00002015y_1y_4y_6 + 1.72\times10^{-6}y_2y_4y_6 \\
&+ 3.958\times10^{-7}y_3y_4y_6 + 1.235\times10^{-7}y_4^2y_6 - 0.00008833y_5y_6 + 0.00002592y_1y_5y_6 - 0.00001742y_2y_5y_6 \\
&+ 1.64\times10^{-8}y_3y_5y_6 - 1.532\times10^{-6}y_4y_5y_6 - 1.02\times10^{-6}y_5^2y_6 - 6.255\times10^{-7}y_6^2 + 8.588\times10^{-7}y_1y_6^2 - 5.559\times10^{-7}y_2y_6^2 \\
&- 5.585\times10^{-9}y_3y_6^2 - 5.575\times10^{-8}y_4y_6^2 - 4.477\times10^{-8}y_5y_6^2 - 5.033\times10^{-1}0y_6^3 + 0.06991y_1y_7 + 0.001902y_1^2y_7 \\
&- 0.04364y_2y_7 - 0.005323y_1y_2y_7 + 0.002109y_2^2y_7 - 0.0005085y_3y_7 + 0.0004046y_1y_3y_7 + 0.001078y_2y_3y_7 \\
&+ 0.0000118y_3^2y_7 - 0.003972y_4y_7 - 0.0006128y_1y_4y_7 + 0.0007717y_2y_4y_7 + 0.00001952y_3y_4y_7 - 0.00001992y_4^2y_7 \\
&- 0.002027y_5y_7 + 0.0001894y_1y_5y_7 + 0.0003707y_2y_5y_7 + 7.84\times10^{-6}y_3y_5y_7 - 6.699\times10^{-8}y_4y_5y_7 - 5.842\times10^{-6}y_5^2y_7 \\
&- 0.00003445y_6y_7 - 2.16\times10^{-7}y_1y_6y_7 + 5.328\times10^{-6}y_2y_6y_7 + 1.706\times10^{-7}y_3y_6y_7 - 1.419\times10^{-8}y_4y_6y_7 \\
&- 7.109\times10^{-7}y_5y_6y_7 - 2.316\times10^{-8}y_6^2y_7 - 0.0007788y_7^2 + 0.00001748y_1y_7^2 + 0.0002561y_2y_7^2 + 9.248\times10^{-6}y_3y_7^2 \\
&- 6.036\times10^{-6}y_4y_7^2 - 6.594\times10^{-7}y_5y_7^2 - 3.222\times10^{-8}y_6y_7^2 + 2.789\times10^{-7}y_7^3 - 0.09614y_1y_8 - 0.006481y_1^2y_8 \\
&- 0.07578y_2y_8 - 0.01064y_1y_2y_8 - 0.01218y_2^2y_8 + 0.004531y_3y_8 + 0.0009064y_1y_3y_8 + 0.004589y_2y_3y_8 \\
&+ 0.00002726y_3^2y_8 - 0.02401y_4y_8 - 0.0004461y_1y_4y_8 + 0.002179y_2y_4y_8 + 0.0001066y_3y_4y_8 - 0.00006922y_4^2y_8 \\
&- 0.007327y_5y_8 + 0.003934y_1y_5y_8 - 0.0005284y_2y_5y_8 + 0.00005616y_3y_5y_8 - 0.00008422y_4y_5y_8 - 0.00008557y_5^2y_8 \\
&- 0.0001124y_6y_8 + 0.0001885y_1y_6y_8 - 0.00008975y_2y_6y_8 + 4.9\times10^{-8}y_3y_6y_8 - 8.671\times10^{-6}y_4y_6y_8 \\
&- 9.134\times10^{-6}y_5y_6y_8 - 1.968\times10^{-7}y_6^2y_8 - 0.0081y_7y_8 + 0.001241y_1y_7y_8 + 0.001322y_2y_7y_8 + 0.00006402y_3y_7y_8 \\
&- 0.00008241y_4y_7y_8 - 0.00006365y_5y_7y_8 - 4.088\times10^{-6}y_6y_7y_8 - 0.00002006y_7^2y_8 - 0.02414y_8^2 - 0.182y_1y_8^2 \\
&+ 0.1261y_2y_8^2 + 0.001514y_3y_8^2 + 0.01061y_4y_8^2 + 0.005221y_5y_8^2 + 0.00007374y_6y_8^2 + 0.004079y_7y_8^2 + 0.02177y_8^3
\end{aligned}
$$

$$x_5 = y_5$$

$$x_6 = y_6$$

$$
\begin{aligned}
x_7 = y_7 &- 0.03866y_1^2 - 0.003413y_1^3 - 0.06335y_1y_2 - 0.006542y_1^2y_2 - 0.1317y_2^2 - 0.00828y_1y_2^2 - 0.01371y_2^3 - 0.1076y_1y_3 \\
&- 0.009591y_1^2y_3 - 0.3346y_2y_3 + 0.03926y_1y_2y_3 + 0.00458y_2^2y_3 + 0.01797y_3^2 - 0.002798y_1y_3^2 - 0.001253y_2y_3^2 \\
&+ 0.000287y_3^3 + 0.236y_1y_4 - 0.0319y_1^2y_4 + 0.003575y_2y_4 - 0.02201y_1y_2y_4 + 0.01759y_2^2y_4 + 0.02132y_3y_4 \\
&- 0.0001487y_1y_3y_4 - 0.007622y_2y_3y_4 + 0.000642y_3^2y_4 + 0.03172y_4^2 + 0.009387y_1y_4^2 + 0.001475y_2y_4^2 \\
&+ 0.000611y_3y_4^2 + 0.0003827y_4^3 - 0.01371y_1y_5 - 0.00366y_1^2y_5 - 0.01301y_2y_5 - 0.00311y_1y_2y_5 - 0.007541y_2^2y_5 \\
&+ 0.001113y_3y_5 - 0.0008713y_1y_3y_5 - 0.000551y_2y_3y_5 - 0.00003594y_3^2y_5 + 0.006997y_4y_5 - 0.0007145y_1y_4y_5 \\
&+ 0.0007672y_2y_4y_5 - 0.00004547y_3y_4y_5 + 0.00001581y_4^2y_5 + 0.007265y_5^2 - 0.0009977y_1y_5^2 + 0.0003618y_2y_5^2 \\
&+ 0.00006402y_3y_5^2 + 0.0001042y_4y_5^2 + 0.00001168y_5^3 + 0.00007289y_1y_6 + 0.00007279y_1^2y_6 - 0.0001411y_2y_6 \\
&+ 0.00003061y_1y_2y_6 + 0.0001349y_2^2y_6 - 0.00003337y_3y_6 + 0.00001517y_1y_3y_6 - 0.00005053y_2y_3y_6 \\
&- 1.574\times10^{-6}y_3^2y_6 + 0.000122y_4y_6 + 0.00005969y_1y_4y_6 + 3.284\times10^{-6}y_2y_4y_6 - 1.979\times10^{-6}y_3y_4y_6 \\
&- 1.703\times10^{-7}y_4^2y_6 + 0.0005265y_5y_6 - 0.0000846y_1y_5y_6 + 0.00001923y_2y_5y_6 + 6.698\times10^{-6}y_3y_5y_6 \\
&+ 0.00001144y_4y_5y_6 + 1.64\times10^{-6}y_5^2y_6 + 0.00001454y_6^2 - 2.6\times10^{-6}y_1y_6^2 + 2.112\times10^{-7}y_2y_6^2 + 2.071\times10^{-7}y_3y_6^2 \\
&+ 3.736\times10^{-7}y_4y_6^2 + 7.497\times10^{-8}y_5y_6^2 + 9.281\times10^{-1}0y_6^3 - 0.1564y_1y_7 + 0.008505y_1^2y_7 - 0.00386y_2y_7 \\
&+ 0.007362y_1y_2y_7 + 0.01324y_3y_7 - 0.003747y_1y_3y_7 - 0.005718y_2y_3y_7 + 0.0004172y_3^2y_7 + 0.02523y_4y_7 \\
&+ 0.004091y_1y_4y_7 - 0.002324y_2y_4y_7 + 0.0006758y_3y_4y_7 + 0.0005056y_4^2y_7 + 0.003422y_5y_7 - 0.001008y_1y_5y_7 \\
&+ 0.00006976y_2y_5y_7 - 0.00002346y_3y_5y_7 + 0.00002762y_4y_5y_7 + 0.00005865y_5^2y_7 + 0.00006843y_6y_7 \\
&+ 7.523\times10^{-6}y_1y_6y_7 - 3.241\times10^{-6}y_2y_6y_7 - 7.901\times10^{-7}y_3y_6y_7 + 6.917\times10^{-7}y_4y_6y_7 + 6.091\times10^{-6}y_5y_6y_7 \\
&+ 1.91\times10^{-7}y_6^2y_7 + 0.006318y_7^2 - 0.0006548y_1y_7^2 - 0.001309y_2y_7^2 + 0.0002191y_3y_7^2 + 0.0002659y_4y_7^2 \\
&+ 0.00001029y_5y_7^2 + 3.709\times10^{-7}y_6y_7^2 + 0.00005266y_7^3 - 0.1212y_1y_8 - 0.008134y_1^2y_8 - 0.1528y_2y_8 \\
&- 0.005201y_1y_2y_8 - 0.03059y_2^2y_8 - 0.0006423y_3y_8 - 0.0136y_1y_3y_8 - 0.01721y_2y_3y_8 + 0.0005924y_3^2y_8 \\
&+ 0.09558y_4y_8 + 0.006536y_1y_4y_8 - 0.002785y_2y_4y_8 + 0.001061y_3y_4y_8 + 0.001373y_4^2y_8 + 0.03784y_5y_8 \\
&- 0.009849y_1y_5y_8 - 0.0006358y_2y_5y_8 + 0.0002865y_3y_5y_8 + 0.0007586y_4y_5y_8 + 0.0002919y_5^2y_8 + 0.00245y_6y_8 \\
&- 0.0003707y_1y_6y_8 - 9.541\times10^{-6}y_2y_6y_8 + 0.00002896y_3y_6y_8 + 0.00006004y_4y_6y_8 + 0.00003159y_5y_6y_8 \\
&+ 8.827\times10^{-7}y_6^2y_8 + 0.05122y_7y_8 - 0.008303y_1y_7y_8 - 0.002967y_2y_7y_8 + 0.0008701y_3y_7y_8 + 0.001638y_4y_7y_8 \\
&+ 0.0004565y_5y_7y_8 + 0.00003247y_6y_7y_8 + 0.0005347y_7^2y_8 - 2.601y_8^2 + 0.3988y_1y_8^2 + 0.005002y_2y_8^2 - 0.03491y_3y_8^2 \\
&- 0.06572y_4y_8^2 - 0.007757y_5y_8^2 - 0.00006561y_6y_8^2 - 0.03226y_7y_8^2 - 0.137y_8^3
\end{aligned}
$$

$$x_8 = y_8 + 0.01041y_1^2 + 0.0009975y_1^3 + 0.0125y_1y_2 + 0.001209y_1^2y_2 + 0.01822y_2^2 + 0.001596y_1y_2^2 + 0.001749y_2^3$$

$$
\begin{aligned}
&- 0.007779y_1y_3 - 0.001094y_1^2y_3 - 0.01174y_2y_3 - 0.00175y_1y_2y_3 + 0.00217y_2^2y_3 - 0.0009694y_3^2 + 0.0001936y_1y_3^2 \\
&+ 0.0002007y_2y_3^2 - 0.00002157y_3^3 + 0.01451y_1y_4 + 0.001721y_1^2y_4 - 0.02039y_2y_4 - 0.002931y_1y_2y_4 - 0.001181y_2^2y_4 \\
&- 0.001725y_3y_4 - 0.00019y_1y_3y_4 + 0.0003721y_2y_3y_4 - 0.0000442y_3^2y_4 - 0.001415y_4^2 - 0.0005517y_1y_4^2 \\
&- 0.0001012y_2y_4^2 - 0.00004194y_3y_4^2 - 0.00002201y_4^3 + 0.002942y_1y_5 - 0.001744y_1^2y_5 - 0.000896y_2y_5 \\
&+ 0.000471y_1y_2y_5 - 0.001578y_2^2y_5 - 0.0002743y_3y_5 + 0.00008367y_1y_3y_5 + 0.00004695y_2y_3y_5 + 2.55\times10^{-6}y_3^2y_5 \\
&- 0.0004617y_4y_5 - 0.00001182y_1y_4y_5 - 0.00005445y_2y_4y_5 + 2.935\times10^{-6}y_3y_4y_5 - 1.973\times10^{-6}y_4^2y_5 - 0.0002851y_5^2 \\
&+ 0.00003543y_1y_5^2 - 0.00004446y_2y_5^2 - 4.7\times10^{-6}y_3y_5^2 - 6.292\times10^{-6}y_4y_5^2 - 9.052\times10^{-7}y_5^3 + 0.00005457y_1y_6 \\
&- 0.00007761y_1^2y_6 + 0.0000285y_2y_6 - 3.839\times10^{-6}y_1y_2y_6 - 0.00007995y_2^2y_6 - 2.901\times10^{-6}y_3y_6 - 7.949\times10^{-7}y_1y_3y_6 \\
&+ 2.447\times10^{-6}y_2y_3y_6 + 1.132\times10^{-7}y_3^2y_6 - 7.802\times10^{-6}y_4y_6 - 3.069\times10^{-6}y_1y_4y_6 + 4.114\times10^{-1}0y_2y_4y_6 \\
&+ 1.242\times10^{-7}y_3y_4y_6 - 1.672\times10^{-8}y_4^2y_6 - 0.00001046y_5y_6 + 2.34\times10^{-6}y_1y_5y_6 - 3.609\times10^{-6}y_2y_5y_6 \\
&- 5.146\times10^{-7}y_3y_5y_6 - 6.581\times10^{-7}y_4y_5y_6 - 1.198\times10^{-7}y_5^2y_6 - 1.158\times10^{-7}y_6^2 + 7.499\times10^{-8}y_1y_6^2 \\
&- 1.076\times10^{-7}y_2y_6^2 - 1.642\times10^{-8}y_3y_6^2 - 2.08\times10^{-8}y_4y_6^2 - 5.06\times10^{-9}y_5y_6^2 + 0.005481y_1y_7 - 0.0001009y_1^2y_7 \\
&- 0.006836y_2y_7 - 0.0001507y_1y_2y_7 + 0.0007276y_2^2y_7 - 0.001083y_3y_7 + 0.0002184y_1y_3y_7 + 0.0003326y_2y_3y_7 \\
&- 0.00003008y_3^2y_7 - 0.001379y_4y_7 - 0.0002434y_1y_4y_7 + 0.0001089y_2y_4y_7 - 0.0000478y_3y_4y_7 - 0.00003177y_4^2y_7 \\
&- 0.0002323y_5y_7 + 0.0000542y_1y_5y_7 - 3.733\times10^{-6}y_2y_5y_7 + 1.641\times10^{-6}y_3y_5y_7 - 2.025\times10^{-6}y_4y_5y_7 \\
&- 3.689\times10^{-6}y_5^2y_7 - 3.601\times10^{-6}y_6y_7 - 5.856\times10^{-7}y_1y_6y_7 + 3.464\times10^{-7}y_2y_6y_7 + 4.978\times10^{-8}y_3y_6y_7 \\
&- 5.114\times10^{-8}y_4y_6y_7 - 3.781\times10^{-7}y_5y_6y_7 - 1.163\times10^{-8}y_6^2y_7 - 0.0003785y_7^2 + 0.00004768y_1y_7^2 + 0.00008062y_2y_7^2 \\
&- 0.00001601y_3y_7^2 - 0.00001734y_4y_7^2 - 6.633\times10^{-7}y_5y_7^2 - 2.408\times10^{-8}y_6y_7^2 - 3.645\times10^{-6}y_7^3 + 0.02835y_1y_8 \\
&- 0.00179y_1^2y_8 + 7.19\times10^{-6}y_2y_8 + 0.001652y_1y_2y_8 - 0.002412y_3y_8 + 0.0007279y_1y_3y_8 + 0.0006777y_2y_3y_8 \\
&- 0.00003753y_3^2y_8 - 0.004693y_4y_8 - 0.0003127y_1y_4y_8 - 0.0002523y_2y_4y_8 - 0.00007737y_3y_4y_8 - 0.00008105y_4^2y_8 \\
&- 0.0008526y_5y_8 + 0.0004095y_1y_5y_8 - 0.0002994y_2y_5y_8 - 0.00002305y_3y_5y_8 - 0.0000446y_4y_5y_8 - 0.00001741y_5^2y_8 \\
&- 0.00001705y_6y_8 + 0.00001257y_1y_6y_8 - 0.0000169y_2y_6y_8 - 2.386\times10^{-6}y_3y_6y_8 - 3.371\times10^{-6}y_4y_6y_8 \\
&- 1.76\times10^{-6}y_5y_6y_8 - 4.321\times10^{-8}y_6^2y_8 - 0.0023y_7y_8 + 0.0004566y_1y_7y_8 + 0.00004006y_2y_7y_8 - 0.00006551y_3y_7y_8 \\
&- 0.0001007y_4y_7y_8 - 0.00002866y_5y_7y_8 - 1.97\times10^{-6}y_6y_7y_8 - 0.00003481y_7^2y_8 - 0.004691y_8^2 - 0.0137y_1y_8^2 \\
&+ 0.01807y_2y_8^2 + 0.002855y_3y_8^2 + 0.003574y_4y_8^2 + 0.0005493y_5y_8^2 + 4.044\times10^{-6}y_6y_8^2 + 0.00193y_7y_8^2 + 0.006093y_8^3.
\end{aligned}
\tag{C.5}
$$

This is the change of coordinates that maps the deterministic system at Hopf point H1 to its inner product normal form. Now using Stratonovich calculus to apply this transformation to the full stochastic series approximation of Eq. (C.4), the resulting transformed 8-dimensional stochastic system is:

$$
\begin{aligned}
dy_1 = &(65.20y_2 - 0.01815y_1^3 - 0.1036y_1^2y_2 - 0.01815y_1y_2^2 - 0.1036y_2^3 - 2.747\times10^{-8}y_6^3 + (4.463 + 0.0406y_1 + 0.03015y_2 \\
&+ 0.004204y_3 - 0.0003061y_4 - 0.002029y_5 - 0.0000902y_6 + 0.002226y_7 + 0.001169y_8)\epsilon)\,dt \\
&+ (4.463 + 0.0406y_1 + 0.001497y_1^2 + 0.03015y_2 + 0.00353y_1y_2 - 0.0008867y_2^2 + 0.004204y_3 + 0.0004247y_1y_3 \\
&- 0.0004167y_2y_3 + 0.00001937y_3^2 - 0.0003061y_4 - 0.001313y_1y_4 + 0.0008467y_2y_4 + 0.00003441y_3y_4 \\
&+ 0.00001588y_4^2 - 0.002029y_5 + 0.0001281y_1y_5 + 0.0004554y_2y_5 - 7.003\times10^{-6}y_3y_5 + 6.999\times10^{-6}y_4y_5 \\
&+ 0.00001385y_5^2 - 0.0000902y_6 + 5.557\times10^{-6}y_1y_6 + 7.505\times10^{-6}y_2y_6 - 2.924\times10^{-7}y_3y_6 + 1.061\times10^{-7}y_4y_6 \\
&+ 1.18\times10^{-6}y_5y_6 + 3.372\times10^{-8}y_6^2 + 0.002226y_7 - 0.0003241y_1y_7 + 0.00008975y_2y_7 + 8.434\times10^{-6}y_3y_7 \\
&+ 0.00007464y_4y_7 + 3.075\times10^{-6}y_5y_7 + 5.729\times10^{-8}y_6y_7 + 9.018\times10^{-6}y_7^2 + 0.001169y_8 - 0.0007381y_1y_8 \\
&+ 0.002781y_2y_8 - 0.00006258y_3y_8 + 0.0002209y_4y_8 + 0.00006919y_5y_8 + 5.14\times10^{-6}y_6y_8 + 0.000112y_7y_8 \\
&- 0.005737y_8^2)\sigma_p \circ dW_2
\end{aligned}
$$

$$
\begin{aligned}
dy_2 = &(-65.20y_1 + 0.1036y_1^3 - 0.01815y_1^2y_2 + 0.1036y_1y_2^2 - 0.01815y_2^3 - 2.461\times10^{-8}y_6^3 + (-0.142 - 0.136y_1 - 0.01293y_2 \\
&- 0.003683y_3 + 0.01428y_4 + 0.007478y_5 + 0.0002041y_6 + 0.002675y_7 + 0.02994y_8)\epsilon)\,dt \\
&+ (-0.142 - 0.136y_1 - 0.006377y_1^2 - 0.01293y_2 + 0.008523y_1y_2 - 0.005854y_2^2 - 0.003683y_3 - 0.0009386y_1y_3 \\
&- 0.002145y_2y_3 + 4.817\times10^{-6}y_3^2 + 0.01428y_4 + 0.001969y_1y_4 - 0.0004839y_2y_4 + 2.419\times10^{-6}y_3y_4 + 0.00005389y_4^2 \\
&+ 0.007478y_5 - 0.00108y_1y_5 + 0.00003628y_2y_5 - 1.373\times10^{-6}y_3y_5 - 0.00003328y_4y_5 - 1.456\times10^{-6}y_5^2 \\
&+ 0.0002041y_6 - 0.00001258y_1y_6 + 7.446\times10^{-6}y_2y_6 - 1.173\times10^{-9}y_3y_6 - 8.76\times10^{-7}y_4y_6 + 6.103\times10^{-7}y_5y_6 \\
&+ 3.297\times10^{-8}y_6^2 + 0.002675y_7 - 0.00008269y_1y_7 - 0.0004312y_2y_7 + 6.291\times10^{-7}y_3y_7 - 0.00001897y_4y_7 \\
&- 3.774\times10^{-7}y_5y_7 + 2.666\times10^{-8}y_6y_7 + 6.855\times10^{-6}y_7^2 + 0.02994y_8 - 0.006372y_1y_8 - 0.001761y_2y_8
\end{aligned}
$$

$$
\begin{aligned}
&+ 2.703{\times}10^{-6} y_3 y_8 - 0.00004385 y_4 y_8 + 0.00004874 y_5 y_8 + 5.753{\times}10^{-6} y_6 y_8 + 0.00004696 y_7 y_8 \\
&- 0.007212 y_8^2)\sigma_p \circ dW_2
\end{aligned}
$$

$$
\begin{aligned}
dy_3 = (&-129 y_3 + 78.75 y_4 - 0.1442 y_1^2 y_3 - 0.1442 y_2^2 y_3 + 0.08727 y_1^2 y_4 + 0.08727 y_2^2 y_4 + (-34.97 + 1.971 y_1 + 6.783 y_2 \\
&- 0.1465 y_3 - 0.07088 y_4 - 0.0517 y_5 - 0.002041 y_6 + 0.1001 y_7 + 0.2672 y_8)\epsilon)\, dt \\
&+ (-34.97 + 1.971 y_1 + 0.3156 y_1^2 + 6.783 y_2 + 0.5874 y_1 y_2 - 1.644 y_2^2 - 0.1465 y_3 + 0.01811 y_1 y_3 - 0.09342 y_2 y_3 \\
&+ 0.00002749 y_3^2 - 0.07088 y_4 - 0.1044 y_1 y_4 - 0.132 y_2 y_4 - 0.000746 y_3 y_4 - 0.002696 y_4^2 - 0.0517 y_5 - 0.0001013 y_1 y_5 \\
&+ 0.0007067 y_2 y_5 - 0.0005354 y_3 y_5 - 0.001172 y_4 y_5 - 9.248{\times}10^{-7} y_5^2 - 0.002041 y_6 - 0.0003917 y_1 y_6 \\
&+ 0.00001553 y_2 y_6 - 0.00001858 y_3 y_6 - 0.00001483 y_4 y_6 + 0.00002339 y_5 y_6 + 1.068{\times}10^{-6} y_6^2 + 0.1001 y_7 \\
&+ 0.01699 y_1 y_7 + 0.03615 y_2 y_7 - 0.0006332 y_3 y_7 - 0.0004557 y_4 y_7 + 0.00007628 y_5 y_7 + 4.315{\times}10^{-6} y_6 y_7 \\
&- 0.00008955 y_7^2 + 0.2672 y_8 - 0.06924 y_1 y_8 - 0.01024 y_2 y_8 - 0.00243 y_3 y_8 - 0.01186 y_4 y_8 - 0.001466 y_5 y_8 \\
&+ 0.0001301 y_6 y_8 - 0.001591 y_7 y_8 - 0.3042 y_8^2)\sigma_p \circ dW_2
\end{aligned}
$$

$$
\begin{aligned}
dy_4 = (&-78.75 y_3 - 129 y_4 - 0.08727 y_1^2 y_3 - 0.08727 y_2^2 y_3 - 0.1442 y_1^2 y_4 - 0.1442 y_2^2 y_4 + (8.565 - 8.088 y_1 + 1.157 y_2 \\
&+ 0.2318 y_3 - 0.2071 y_4 + 0.1645 y_5 - 0.0009561 y_6 - 0.2449 y_7 + 1.043 y_8)\epsilon)\, dt \\
&+ (8.565 - 8.088 y_1 - 0.9236 y_1^2 + 1.157 y_2 + 1.931 y_1 y_2 - 0.3523 y_2^2 + 0.2318 y_3 + 0.2578 y_1 y_3 - 0.02374 y_2 y_3 \\
&- 0.001081 y_3^2 - 0.2071 y_4 + 0.1862 y_1 y_4 + 0.07097 y_2 y_4 + 0.001607 y_3 y_4 - 0.002156 y_4^2 + 0.1645 y_5 - 0.002241 y_1 y_5 \\
&+ 0.01171 y_2 y_5 + 0.0009797 y_3 y_5 - 0.0008292 y_4 y_5 - 0.0009024 y_5^2 - 0.0009561 y_6 - 0.0002575 y_1 y_6 - 0.0003167 y_2 y_6 \\
&+ 0.00002432 y_3 y_6 - 0.00002028 y_4 y_6 - 0.0000933 y_5 y_6 - 3.125{\times}10^{-6} y_6^2 - 0.2449 y_7 - 0.03275 y_1 y_7 + 0.03251 y_2 y_7 \\
&+ 0.0004389 y_3 y_7 - 0.001258 y_4 y_7 + 0.00004382 y_5 y_7 - 2.19{\times}10^{-6} y_6 y_7 - 0.0002962 y_7^2 + 1.043 y_8 - 0.02192 y_1 y_8 \\
&+ 0.02352 y_2 y_8 + 0.01052 y_3 y_8 - 0.008399 y_4 y_8 - 0.009548 y_5 y_8 - 0.0006858 y_6 y_8 - 0.003703 y_7 y_8 \\
&+ 0.6466 y_8^2)\sigma_p \circ dW_2
\end{aligned}
$$

$$
dy_5 = (-100 y_5 + y_6)\, dt
$$

$$
dy_6 = -100 y_6\, dt
$$

$$
\begin{aligned}
dy_7 = (&-161.9 y_7 + 0.2186 y_1^2 y_7 + 0.2186 y_2^2 y_7 + (66.82 + 4.296 y_1 - 12.11 y_2 + 0.6183 y_3 - 1.981 y_4 + 0.02958 y_5 \\
&+ 0.007109 y_6 + 0.3975 y_7 - 33.98 y_8)\epsilon)\, dt \\
&+ (66.82 + 4.296 y_1 + 2.071 y_1^2 - 12.11 y_2 + 0.2461 y_1 y_2 + 2.106 y_2^2 + 0.6183 y_3 + 0.1022 y_1 y_3 - 0.06549 y_2 y_3 \\
&+ 0.003868 y_3^2 - 1.981 y_4 + 0.02738 y_1 y_4 - 0.01803 y_2 y_4 + 0.001115 y_3 y_4 + 0.01241 y_4^2 + 0.02958 y_5 + 0.04209 y_1 y_5 \\
&- 0.02005 y_2 y_5 + 0.0001382 y_3 y_5 + 0.005125 y_4 y_5 + 0.002817 y_5^2 + 0.007109 y_6 + 0.000249 y_1 y_6 - 0.001864 y_2 y_6 \\
&- 0.00005915 y_3 y_6 + 0.0001843 y_4 y_6 + 0.0002819 y_5 y_6 + 7.961{\times}10^{-6} y_6^2 + 0.3975 y_7 + 0.09244 y_1 y_7 + 0.009339 y_2 y_7 \\
&- 0.0004824 y_3 y_7 + 0.007767 y_4 y_7 + 0.001874 y_5 y_7 + 0.00006757 y_6 y_7 + 0.001589 y_7^2 - 33.98 y_8 - 3.851 y_1 y_8 \\
&+ 1.047 y_2 y_8 + 0.2007 y_3 y_8 - 0.1764 y_4 y_8 - 0.02121 y_5 y_8 + 0.0006618 y_6 y_8 - 0.06044 y_7 y_8 - 1.396 y_8^2)\sigma_p \circ dW_2
\end{aligned}
$$

$$
\begin{aligned}
dy_8 = (&-80.14 y_8 + 1.163{\times}10^{-8} y_6^3 + 0.1424 y_1^2 y_8 + 0.1424 y_2^2 y_8 + (-5.811 - 0.6889 y_1 + 0.1704 y_2 + 0.03839 y_3 \\
&- 0.03887 y_4 - 0.008328 y_5 - 0.0001326 y_6 - 0.01428 y_7 - 0.07152 y_8)\epsilon)\, dt \\
&+ (-5.811 - 0.6889 y_1 + 0.05968 y_1^2 + 0.1704 y_2 - 0.03766 y_1 y_2 + 0.0515 y_2^2 + 0.03839 y_3 - 0.003114 y_1 y_3 \\
&+ 0.01083 y_2 y_3 - 0.00005973 y_3^2 - 0.03887 y_4 - 0.004438 y_1 y_4 - 0.002003 y_2 y_4 + 3.153{\times}10^{-6} y_3 y_4 - 0.0001274 y_4^2 \\
&- 0.008328 y_5 + 0.01163 y_1 y_5 + 0.0003977 y_2 y_5 + 0.0000379 y_3 y_5 + 0.00002289 y_4 y_5 - 0.00005534 y_5^2 - 0.0001326 y_6 \\
&+ 0.0007129 y_1 y_6 - 0.00005582 y_2 y_6 + 1.163{\times}10^{-6} y_3 y_6 + 5.722{\times}10^{-7} y_4 y_6 - 6.295{\times}10^{-6} y_5 y_6 - 2.112{\times}10^{-7} y_6^2 \\
&- 0.01428 y_7 + 0.000859 y_1 y_7 - 0.0002826 y_2 y_7 + 0.00002256 y_3 y_7 - 0.0001627 y_4 y_7 - 0.00001886 y_5 y_7 \\
&- 5.813{\times}10^{-7} y_6 y_7 - 0.00003916 y_7^2 - 0.07152 y_8 - 0.001575 y_1 y_8 + 0.01153 y_2 y_8 + 0.00054 y_3 y_8 - 0.000912 y_4 y_8 \\
&- 0.0006741 y_5 y_8 - 0.00004171 y_6 y_8 - 0.0005302 y_7 y_8 + 0.03689 y_8^2)\sigma_p \circ dW_2. \qquad \text{(C.6)}
\end{aligned}
$$

C.2 Reduction and averaging of Fokker-Planck operator

Step 3 (approximate reduction to 2D and transformation to polar coordinates), then gives:

$$\begin{aligned}
dr &= (-0.01815r^3 + \epsilon(4.463\cos\theta - 0.1420\sin\theta) + \epsilon r(0.01384 + 0.02677\cos 2\theta - 0.05294\sin 2\theta))\,dt \\
&\quad + (4.463\cos\theta - 0.1420\sin\theta + r(0.01384 + 0.02677\cos 2\theta - 0.05294\sin 2\theta) + r^2(0.003032\cos\theta - 0.005102\sin\theta \\
&\quad - 0.001535\cos 3\theta + 0.0007518\sin 3\theta))\sigma_p \circ dW_2 \\
d\theta &= (-65.2 + 0.1036r^2 + \epsilon((-0.1420\cos\theta - 4.463\sin\theta)/r - 0.08309 - 0.05294\cos 2\theta - 0.02677\sin 2\theta))\,dt \\
&\quad + ((-0.1420\cos\theta - 4.463\sin\theta)/r - 0.08309 - 0.05294\cos 2\theta - 0.02677\sin 2\theta + r(-0.007129\cos\theta + 0.002422\sin\theta \\
&\quad + 0.0007518\cos 3\theta + 0.001535\sin 3\theta))\sigma_p \circ dW_2.
\end{aligned} \tag{C.7}$$

Step 4 (averaging the corresponding Fokker-Planck equation around the cycle and selecting a weakly equivalent SDE by Cholesky decomposition) gives:

$$\begin{aligned}
dr &= (-0.01815r^3 + (0.01384\epsilon + 0.002875\sigma_p^2)r + 4.984\sigma_p^2\frac{1}{r})\,dt \;+\; (3.157 + 0.002567r^2)\sigma_p \circ dW_A \\
d\theta &= (-65.20 - 0.08309\epsilon - 0.01059\sigma_p^2 + 0.1036r^2)\,dt - 0.001919r\sigma_p \circ dW_A + (3.157\frac{1}{r} - 0.000179r)\sigma_p \circ dW_B.
\end{aligned} \tag{C.8}$$

Finally, linearly rescaling the radius by $R = \sqrt{0.01815}\,r$,

$$\begin{aligned}
dR &= (-R^3 + (0.01384\epsilon + 0.002875\sigma^2)R + 0.09045\sigma^2\frac{1}{R})\,dt \;+\; (0.4253 + 0.01905R^2)\sigma \circ dW_A \\
d\theta &= (-65.20 - 0.08309\epsilon - 0.01059\sigma^2 + 5.706R^2)\,dt - 0.01424R\sigma \circ dW_A + (0.4253\frac{1}{R} - 0.001328R)\sigma \circ dW_B.
\end{aligned} \tag{C.9}$$

The results for bifurcation points H2 and H3 were computed similarly. In the case of H2, where both pyramidal and spiny stellate cell populations receive noise input, the results were found for both correlated and uncorrelated noise to the two populations and there was no significant difference in the resulting parameters.

The results using an ϵ-dependent normal form transformation derived from the 9-dimensional extended systems were also found in each case, to include third-order correction terms that show whether the key parameters vary significantly through the bifurcation. Those correction terms were negligible in each case (coefficients $< 10^{-3}$), so the simpler equations were preferred.

C.3 Simulation verifies oscillation properties preserved

Figures C.1 - C.3 on the next pages show the results of simulation comparing statistics in the full Jansen-Rit system near bifurcation point H1 and the transformed and averaged systems. This shows that oscillation statistics on a longer time scale than the period are indeed preserved to a good approximation, despite the extreme numerical ill-conditioning of the Jansen-Rit system.

The power spectral density of the Hilbert amplitude time series is approximately preserved for frequencies below the dominant oscillation at $\sim 10\,\text{Hz}$. For the original Jansen-Rit model the power spectrum of the amplitude time series shows a wide peak near 10 Hz, while the reduced normal form system before averaging does not. The cause of this discrepancy has not yet been diagnosed and it may be an artifact of incorrect numerical treatment of the Hilbert transform and power spectral estimation. It may be due to the bistability of the original system near point H1: during the 100 s simulation 4.9% of the sample paths switched from the limit cycle to the stable node described in Figure 2.2. Finally, it may be a difference between amplitude dynamics of the original and transformed systems due to the normal form system having oscillations that are closer to harmonic. To investigate this issue the future work will be to examine the power spectrum of the intermediate 8-dimensional system after linear transformation but before near-identity transformation.

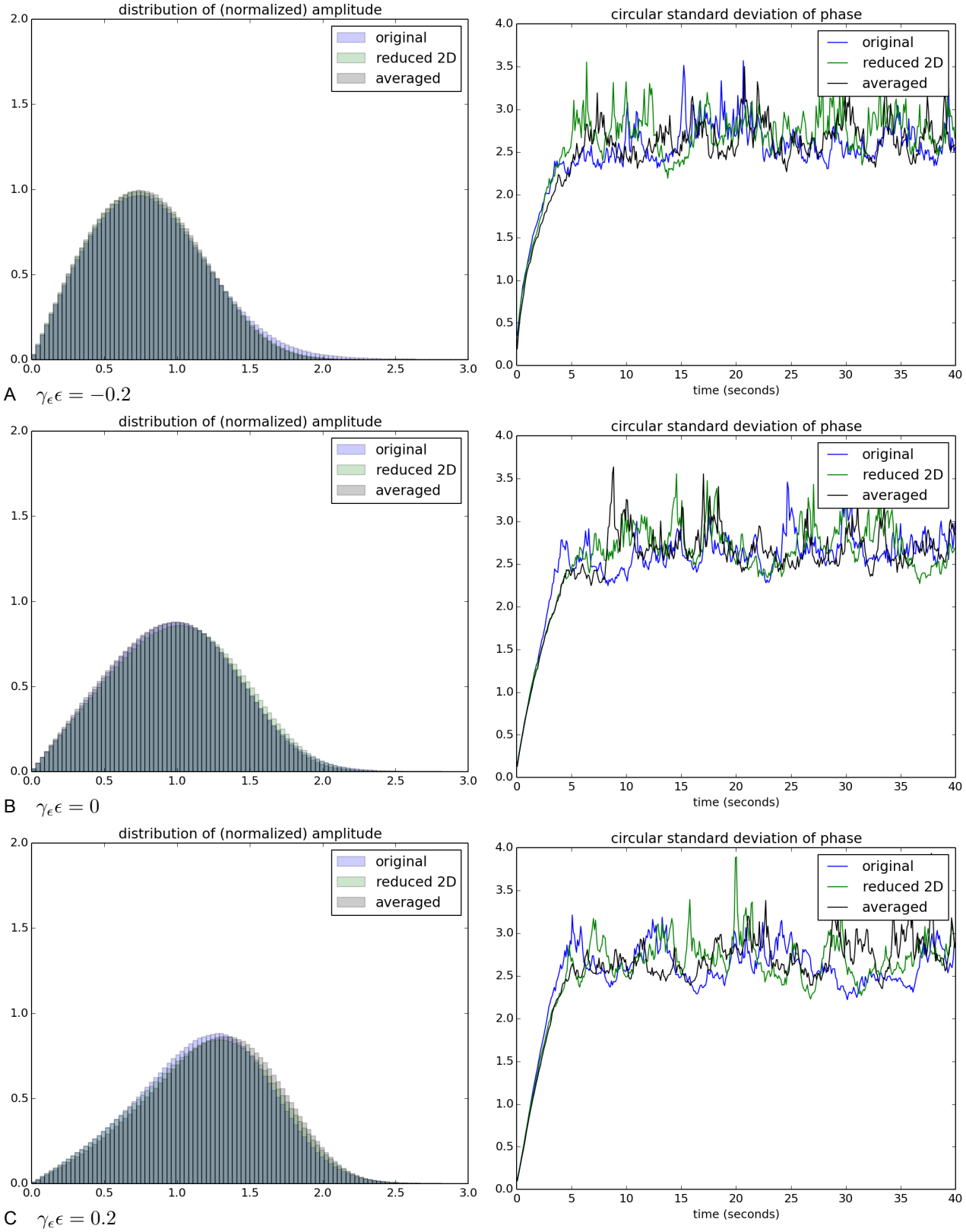
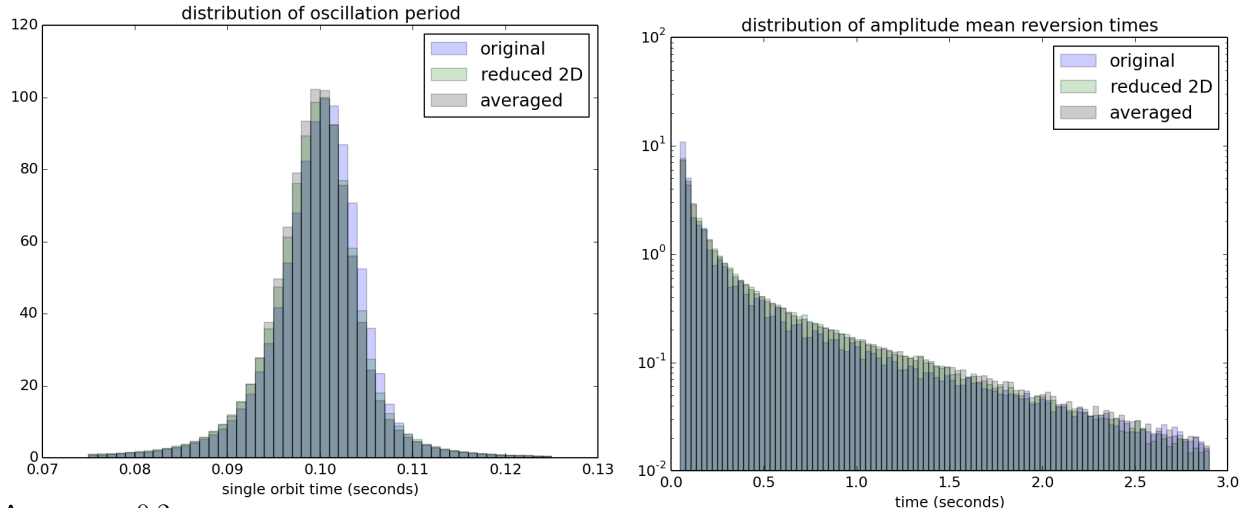
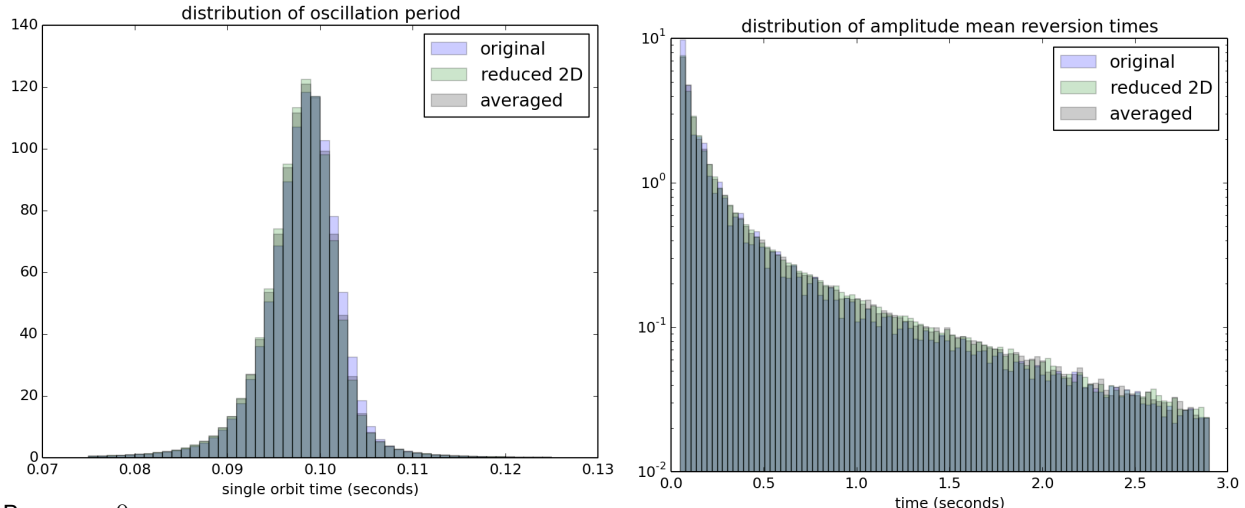


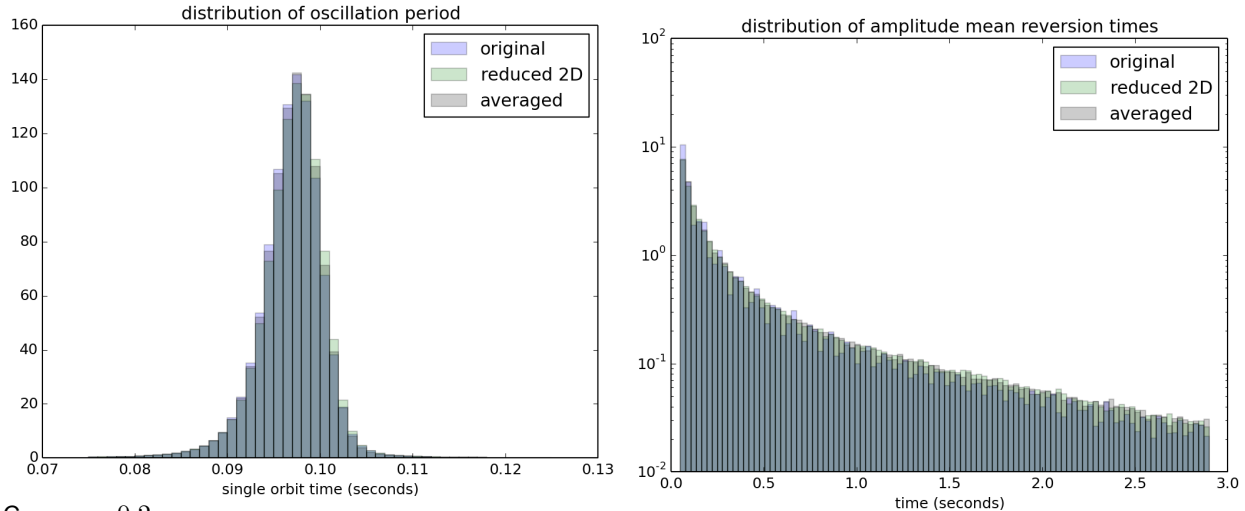
Figure C.1: Results for Jansen-Rit system at supercritical Hopf point H1: Normalized amplitude distribution (left) and phase diffusion (right) before, at and after the deterministic bifurcation point $\epsilon = 0$, superimposing the original 8D system (blue), 2D transformed system (green) and final averaged system (black).



A $\gamma_\epsilon \epsilon = -0.2$



B $\gamma_\epsilon \epsilon = 0$



C $\gamma_\epsilon \epsilon = 0.2$

Figure C.2: Results for Jansen-Rit system at supercritical Hopf point H1: Distribution of stochastic period (left) and mean reversion time (right) before, at and after the deterministic bifurcation point $\epsilon = 0$, superimposing the original 8D system (blue), 2D transformed system (green) and final averaged system (black).

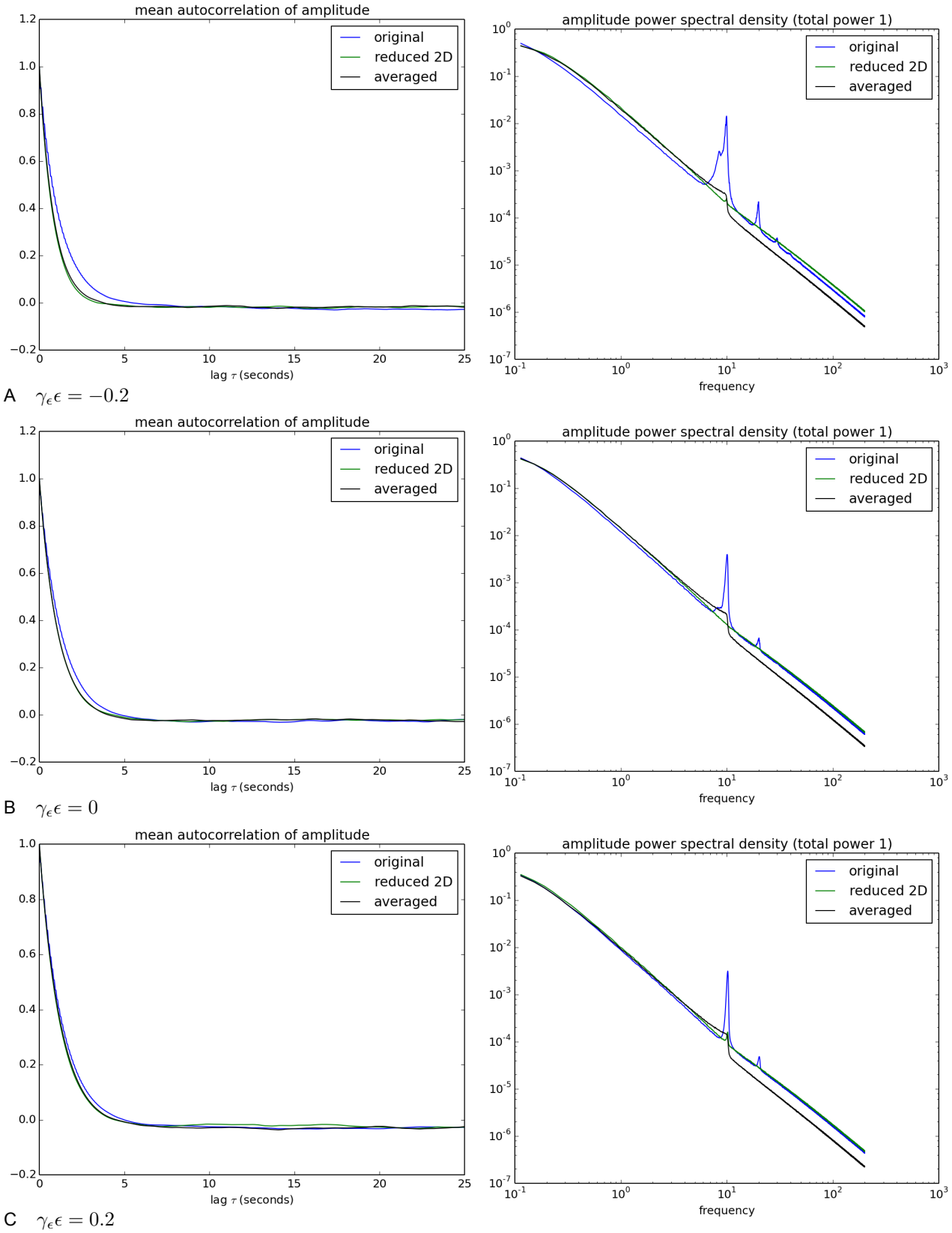


Figure C.3: Results for Jansen-Rit system at point H1: autocorrelation of amplitude (left) and power spectral density (right) before, at and after the deterministic bifurcation point $\epsilon = 0$, superimposing original 8D system (blue), 2D transformed system (green) and averaged system (black).

Appendix D

A NormalForm package for Mathematica

To use computer algebra for automatic derivation of local normal forms (of semisimple and inner-product style) for arbitrary systems, a Mathematica package was created. This package can be given as input an arbitrary dynamical system of ordinary differential equations. It will then explicitly compute the smooth transformation that maps the dynamical system (near a local bifurcation of equilibria) to a simpler low-dimensional normal form system that preserves the same local behavior.

For stochastic dynamical systems near a Hopf bifurcation, this `NormalForm` package also implements the pipeline of transformations to reduce the local system near bifurcation and output the 2-dimensional weak approximation standard form developed in Chapter 4.

The Mathematica code has been released as a free software project under a GPL license and is available at https://github.com/mattja/NormalForm.

This appendix explains the usage of the package.

D.1 Normal forms near local bifurcations of equilibria

`NormalFormTransformation[rhs, {x1,...,xn}, {u1,...,un}, m]` transforms the dynamical system with right hand side `rhs` (expressed in original variables `{xi}`) to a simpler system (normal form to order `m`) in the new variables `{ui}`. It returns a pair `{newrhs, trans}` where `newrhs` is

the transformed system and `trans` is a smooth invertible coordinate transformation that maps `rhs` to `newrhs`. N.B. it is assumed that the linear part of the system has already been transformed to Jordan real form.

Options:

`Verbose->True` will cause it to print out working at each step.

`BifurcationParameters->{eps1, eps2, ...}` set which symbols in `rhs` should be interpreted as bifurcation parameters. Default is `Global`\[Epsilon]`.

`AsymptoticScaling->{symbol1^ exponent1,...}` advise what asymptotic scaling to assume when truncating the resulting power series. The default is `{x1,...,xn,Sqrt[\[Epsilon]]}` (which means that ϵ is taken to be the same order as the x_i squared).

`Extended`: whether to compute the normal form of the extended system, that is with phase space extended with dimensions for the (rescaled) bifurcation parameters and their equations $\dot{\epsilon} = 0$, deriving a transformation dependent on the bifurcation parameters. If False (the default) then the normal form will be found with respect to the dynamical variables only.

D.2 Transformation of stochastic Hopf bifurcation

```
TransformNoisyHopf[rhs, {x1,...,xn}, {\[Sigma]1,...,\[Sigma]n},
                   {\[Xi]1,...\[Xi]n}, r, {new\[Xi]1, new\[Xi]2}]
```

takes the stochastic dynamical system with right hand side `rhs` (expressed in variables `{xi}`, small noise parameters `{\[Sigma]i}` and Langevin noise symbols `{\[Xi]i}` with Stratonovich interpretation of any multiplicative noise) and transforms it to a simple circular 2 dimensional Hopf normal form system (expressed in new polar variables `{r, \[Theta]}` and new Langevin noise symbols `{new\[Xi]1, new\[Xi]2}`).

N.B. It is assumed that the linear part of the system has already been transformed to Jordan real form, with Hopf bifurcation in first two variables at the origin.

Options:

`Verbose->True` (as above)

`BifurcationParameters->{eps}` (as above)

`AsymptoticScaling->{symbol1^ exponent1,...}`

with default value `{x1,...,xn,Sqrt[\[Epsilon],\[Sigma]1,...\[Sigma]n}` (i.e. by default the noise strengths σ_i are taken to be of the same order as the x_i when truncating the resulting power series)

`MaxOrder->n` The default value 3 (compute all terms, including noise effects, up to third order) is sensible, unless the Hopf is degenerate.

`Extended`: whether to compute the normal form of the extended system, that is with phase space extended with dimensions for the (rescaled) bifurcation parameters and their equations $\dot{\epsilon} = 0$, deriving a transformation dependent on the bifurcation parameters. If False (the default) then the normal form will be found with respect to the dynamical variables only.

`Average`: whether to average around the cycle. (default True)

`Rescale`: whether to linearly rescale the radial variable to make the coefficient of the R^3 term -1. (default True)

D.3 Asymptotic tools

`MultiSeries[vectorField, {x1^ exp1, ...}, maxOrder]` Implements computation with truncated multivariate power series, supporting arbitrary asymptotic scaling of variables.

`TransformContravariant[U, R]` applies the near-identity coordinate transformation `U` to transform the contravariant vector field `R(u)`. Both `U` and `R` should be given in the form of `MultiSeries`.

`BalanceMatrix[A]` returns the pair `{T, B}` where `T` is a similarity transformation and `B` is the transformed matrix, $B = T^{-1} A T$, such that B is as close to symmetric as possible. This is used to improve an ill-conditioned matrix `A`, allowing eigenvalues and eigenvectors to be computed more precisely from matrix `B` (Parlett and Reinsch, 1969).

It remains to automate the preparation step (translation and linear transformation to put the linear part of system in Jordan real form with Hopf at the origin in the first two variables). Currently you need to do this preparation separately before using `NormalFormTranformation[]` or `TransformNoisyHopf[]`. In future it will be useful to implement the Lie algebra based normal form algorithms given in Murdock (2003), which besides being more memory efficient also give explicitly the inverse transformation.

Appendix E

Large scale numerical simulation of stochastic differential equations

To perform numerical integration of SDE systems using integration algorithms of strong convergence, and satisfying our need to support multivariate Ito and Stratonovich equations with multiple driving Wiener processes, a prototype python package `sdeint` was written. This implements the Ito and Stratonovich stochastic Runge-Kutta schemes SRI2 and SRS2 of Rößler (2010), selected simpler algorithms of Maruyama (1955), Rümelin (1982), and Kloeden and Platen (1999). We use the algorithms for approximation of multiple stochastic integrals of Kloeden and Platen (1992) and Wiktorsson (2001).

Then, to make it easy to perform massively parallel simulation of the SDE systems on a high performance computing cluster or cloud service that has separate compute nodes separated by network links, the python package `nsim` was written. This package internally uses `sdeint` on each computing node to perform the SDE integrations. It then allows to analyze the resulting time series interactively while keeping the data in distributed arrays spread over the cluster.

Both packages have been released as free software projects under a GPL license and can be downloaded from https://github.com/mattja.

This appendix describes the usage of these packages.

E.1 Numerical integration of SDE

`sdeint` is a collection of numerical algorithms for integrating Ito and Stratonovich stochastic ordinary differential equations (SODEs). It has simple functions that can be used in a similar way to `scipy.integrate.odeint` or MATLAB's `ode45`.

Functions

`itoint(f, G, y0, tspan)` for Ito equation $\mathbf{dy} = \mathbf{f}(\mathbf{y}, t)dt + G(\mathbf{y}, t)\mathbf{dW}$

`stratint(f, G, y0, tspan)` for Stratonovich equation $\mathbf{dy} = \mathbf{f}(\mathbf{y}, t)dt + G(\mathbf{y}, t) \circ \mathbf{dW}$

These work with scalar or vector equations. They will choose an algorithm for you. Or you can use a specific algorithm directly:

Specific algorithms

`itoEuler(f, G, y0, tspan)`: the Euler-Maruyama algorithm for Ito equations.

`stratHeun(f, G, y0, tspan)`: the Stratonovich Heun algorithm for Stratonovich equations.

`itoSRI2(f, G, y0, tspan)`: the Rößler (2010) order 1.0 strong Stochastic Runge-Kutta algorithm SRI2 for Ito equations.

`itoSRI2(f, [g1,...,gm], y0, tspan)`: as above, with G matrix given as a separate function for each column (gives speedup for large m or complicated G).

`stratSRS2(f, G, y0, tspan)`: the Rößler (2010) order 1.0 strong Stochastic Runge-Kutta algorithm SRS2 for Stratonovich equations.

`stratSRS2(f, [g1,...,gm], y0, tspan)`: as above, with G matrix given as a separate function for each column (gives speedup for large m or complicated G).

`stratKP2iS(f, G, y0, tspan)`: the Kloeden and Platen two-step implicit order 1.0 strong algorithm for Stratonovich equations.

For more information and advanced options see the documentation for each function.

Utility functions

`deltaW(N, m, h)`: Generate increments of m independent Wiener processes for each of N time intervals of length h.

Repeated integrals by the method of Kloeden and Platen (1992):

`Ikpw(dW, h, n=5)`: Approximate repeated Ito integrals.

`Jkpw(dW, h, n=5)`: Approximate repeated Stratonovich integrals.

Repeated integrals by the method of Wiktorsson (2001):

`Iwik(dW, h, n=5)`: Approximate repeated Ito integrals.

`Jwik(dW, h, n=5)`: Approximate repeated Stratonovich integrals.

Examples

Integrate the one-dimensional Ito equation $dx = -(a + xb^2)(1 - x^2)\,dt + b(1 - x^2)\,dW$,
with initial condition $x_0 = 0.1$.

```
import numpy as np
import sdeint

a = 1.0
b = 0.8
tspan = np.linspace(0.0, 5.0, 5001)
x0 = 0.1

def f(x, t):
    return -(a + x*b**2)*(1 - x**2)

def g(x, t):
    return b*(1 - x**2)

result = sdeint.itoint(f, g, x0, tspan)
```

Integrate the two-dimensional vector Ito equation $\mathbf{dx} = A.\mathbf{x}\,dt + B.\mathbf{dW}$,
where `x = (x1, x2)`, `dW = (dW1, dW2)` and with initial condition $\mathbf{x}_0 = (3.0, 3.0)$.

```
import numpy as np
import sdeint

A = np.array([[-0.5, -2.0],
              [ 2.0, -1.0]])

B = np.diag([0.5, 0.5]) # diagonal, so independent driving Wiener processes

tspan = np.linspace(0.0, 10.0, 10001)
x0 = np.array([3.0, 3.0])

def f(x, t):
    return A.dot(x)

def G(x, t):
    return B

result = sdeint.itoint(f, G, x0, tspan)
```

E.2 Distributed computation and time series analysis

`nsim` is for systems in physics, biology and finance that are modelled in continuous time with differential equations. nsim makes it easy to define and simulate these (including proper treatment of noise) and to analyze the resulting time series.

Automatic parallel computing / cluster computing: For multiple or repeated simulations, nsim distributes these across a cluster or Amazon EC2 cloud (or across the CPUs of one computer) without needing to do any parallel programming. It is first necessary to configure an IPython cluster, for example on a single computer by typing `ipcluster start`)

To define a scalar or vector ODE system, subclass `ODEModel` (see examples). To define a scalar or vector SDE system, subclass `ItoModel` or `StratonovichModel`. Multiple driving Wiener processes are now supported. Order 1.0 strong stochastic Runge-Kutta algorithms (Rößler, 2010) are used for SDE integration by default.

Model parameters can be specified as random distributions, to create multiple non-identical simulations.

The `NetworkModel` class allows you to simulate many subsystems coupled together into a network, with the network structure specified as a weighted directed graph. Sub-models can all be identical but they don't have to be. (The networkx package can optionally be used to generate various kinds of random, clustered and small world graphs useful in a NetworkModel). The sub-models in a NetworkModel can even be other NetworkModels, for simulating networks of networks.

Analyzing time series

Besides time series from simulations, empirical time series data can also be loaded from MATLAB .mat files or .EDF files for distributed analysis.

nsim provides a `Timeseries` class. This is a numpy array. It allows slicing the array by time instead of by array index, e.g. can write `ts.t[10.5:30]` to slice from t=10.5 to t=30 seconds. When manipulating the array it will keep track of any channel names (or variable names) of a multivariate time series.

As well as the usual methods of numpy arrays, the `Timeseries` objects have extra methods for easy filtering, plotting and analysis. Analyses can be chained together in a pipeline. For example with a `Timeseries` instance `ts` you can write a chain of analyses like `ts.t[10:30].bandpass(20, 35).hilbert().abs().plot()` This can be extended with your own analysis functions by calling `Timeseries.add_analyses()` Analysis of multiple time series is distributed on the cluster, without needing to do any parallel programming.

Bibliography

Aburn, M. J., Holmes, C. A., Roberts, J. A., Boonstra, T. W., and Breakspear, M. (2012). Critical fluctuations in cortical models near instability. *Frontiers in Physiology* 3.331.

Andronov, A. A. and Witt, A. (1930). Sur la théorie mathématique des autooscillations. *CR Acad. Sci. Paris* 190, 256–258.

Anteneodo, C. and Riera, R. (2005). Additive-multiplicative stochastic models of financial mean-reverting processes. *Physical Review E* 72, 026106.

Arnold, L. and Imkeller, P. (1998). Normal forms for stochastic differential equations. *Probability Theory and Related Fields* 110, 559–588.

Aronson, D. G., Ermentrout, G. B., and Kopell, N. (1990). Amplitude response of coupled oscillators. *Physica D: Nonlinear Phenomena* 41, 403–449.

Ashwin, P. and Rodrigues, A. (2016). Hopf normal form with symmetry and reduction to systems of nonlinearly coupled phase oscillators. *Physica D: Nonlinear Phenomena* 325, 14–24.

Assisi, C., Jirsa, V., and Kelso, J. (2005). Synchrony and clustering in heterogeneous networks with global coupling and parameter dispersion. *Physical review letters* 94, 18106.

Averbeck, B. B., Latham, P. E., and Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Reviews Neuroscience* 7, 358–366.

Babajani-Feremi, A. and Soltanian-Zadeh, H. (2010). Multi-area neural mass modeling of EEG and MEG signals. *NeuroImage*. Computational Models of the Brain 52, 793–811.

Baker, A. (2013). Simplicity. *The Stanford Encyclopedia of Philosophy*. Ed. by E. N. Zalta. Fall 2013.

Bastos, A. M., Vezoli, J., and Fries, P. (2015). Communication through coherence with inter-areal delays. *Current Opinion in Neurobiology*. SI: Brain rhythms and dynamic coordination 31, 173–180.

Bedard, C., Kroeger, H., and Destexhe, A. (2006). Does the 1/f frequency scaling of brain signals reflect self-organized critical states? *Physical review letters* 97, 118102.

Benjamin, O., Fitzgerald, T. H., Ashwin, P., Tsaneva-Atanasova, K., Chowdhury, F., Richardson, M. P., and Terry, J. R. (2012). A phenomenological model of seizure initiation suggests network structure may explain seizure frequency in idiopathic generalised epilepsy. *The Journal of Mathematical Neuroscience* 2, 1.

Berger, H. (1929). Über das Elektrenkephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten* 87, 527–570.

Binney, J. (1992). *The Theory of Critical Phenomena: An Introduction to the Renormalization Group*. Clarendon Press.

Blenkinsop, A., Valentin, A., Richardson, M. P., and Terry, J. R. (2012). The dynamic evolution of focal-onset epilepsies–combining theoretical and clinical observations. *European Journal of Neuroscience* 36, 2188–2200.

Bogoliubov, N. N. and Mitropolsky, Y. A. (1961). *Asymptotic Methods in the Theory of Non-linear Oscillations*. CRC Press.

Boly, M., Garrido, M. I., Gosseries, O., Bruno, M., Boveroux, P., Schnakers, C., Massimini, M., Litvak, V., Laureys, S., and Friston, K. (2011). Preserved Feedforward But Impaired Top-Down Processes in the Vegetative State. *Science* 332, 858–862.

Boonstra, T. (2007). Neural dynamics of human motor control. PhD thesis. Amsterdam: Vrije Universiteit.

Boonstra, T., Daffertshofer, A., Roerdink, M., Flipse, I., Groenewoud, K., and Beek, P. (2009a). Bilateral motor unit synchronization of leg muscles during a simple dynamic balance task. *European Journal of Neuroscience* 29, 613–622.

Boonstra, T., van Wijk, B., Praamstra, P., and Daffertshofer, A. (2009b). Corticomuscular and bilateral EMG coherence reflect distinct aspects of neural synchronization. *Neuroscience letters* 463, 17–21.

Braun, J. and Mattia, M. (2010). Attractors and noise: Twin drivers of decisions and multistability. *Neuroimage* 52, 740–751.

Breakspear, M., Heitmann, S., and Daffertshofer, A. (2010). Generative Models of Cortical Oscillations: Neurobiological Implications of the Kuramoto Model. *Frontiers in Human Neuroscience* 4, 1–14.

Breakspear, M., Roberts, J. A., Terry, J. R., Rodrigues, S., Mahant, N., and Robinson, P. A. (2006). A unifying explanation of primary generalized seizures through nonlinear brain modeling and bifurcation analysis. *Cerebral Cortex* 16, 1296.

Breakspear, M. and Terry, J. R. (2002). Detection and description of non-linear interdependence in normal multichannel human EEG data. *Clinical Neurophysiology* 113, 735–753.

Breakspear, M. and Jirsa, V. K. (2007). Neuronal Dynamics and Brain Connectivity. *Handbook of Brain Connectivity*. Ed. by V. K. Jirsa and A. R. McIntosh. Understanding Complex Systems. Springer Berlin Heidelberg, pp. 3–64.

Breakspear, M., Terry, J. R., and Friston, K. J. (2003). Modulation of excitatory synaptic coupling facilitates synchronization and complex dynamics in a biophysical model of neuronal dynamics. *Network: Computation in Neural Systems* 14, 703–732.

Brown, E., Moehlis, J., and Holmes, P. (2004). On the phase reduction and response dynamics of neural oscillator populations. *Neural Computation* 16, 673–715.

Brunel, N., Chance, F. S., Fourcaud, N., and Abbott, L. F. (2001). Effects of Synaptic Noise and Filtering on the Frequency Response of Spiking Neurons. *Physical Review Letters* 86, 2186–2189.

Brunel, N. and Wang, X. (2001). Effects of Neuromodulation in a Cortical Network Model of Object Working Memory Dominated by Recurrent Inhibition. *Journal of Computational Neuroscience* 11, 63–85.

Bruns, A. (2004). Fourier-, Hilbert- and wavelet-based signal analysis: are they really different approaches? *Journal of Neuroscience Methods* 137, 321–332.

Bruns, A., Eckhorn, R., Jokeit, H., and Ebner, A. (2000). Amplitude envelope correlation detects coupling among incoherent brain signals. *NeuroReport* 11, 1509–1514.

Buice, M. A., Cowan, J. D., and Chow, C. C. (2010). Systematic Fluctuation Expansion for Neural Network Activity Equations. *Neural Computation* 22, 377–426.

Burrage, K., Burrage, P. M., and Tian, T. (2004). Numerical Methods for Strong Solutions of Stochastic Differential Equations: An Overview. *Proceedings: Mathematical, Physical and Engineering Sciences* 460, 373–402.

Buzsáki, G. (2006). *Rhythms of the Brain*. Oxford University Press.

Buzsáki, G. and Draguhn, A. (2004). Neuronal Oscillations in Cortical Networks. *Science* 304, 1926–1929.

Cai, D., Tao, L., Shelley, M., and McLaughlin, D. W. (2004). An effective kinetic representation of fluctuation-driven neuronal networks with application to simple and complex cells in visual cortex. *Proceedings of the National Academy of Sciences of the United States of America* 101, 7757–7762.

Campbell, S. (2007). Time delays in neural systems. *Handbook of Brain Connectivity*. Springer, pp. 65–90.

Chatterjee, A. (2009). Identification of System Nonlinearity Structure by Harmonic Probing. *Proceedings of the IMAC-XXVII*.

Chávez, M., Martinerie, J., and Le Van Quyen, M. (2003). Statistical assessment of nonlinear causality: application to epileptic EEG signals. *Journal of Neuroscience Methods* 124, 113–128.

Chialvo, D. R. (2010). Emergent complex neural dynamics. *Nature Physics* 6, 744–750.

Chialvo, D. R. (2014). Critical Brain Dynamics at Large Scale. *Criticality in Neural Systems*. Ed. by S. I. D. Plenz and E. Niebur. Wiley-VCH, pp. 43–66.

Chowdhury, F. A., Woldman, W., FitzGerald, T. H. B., Elwes, R. D. C., Nashef, L., Terry, J. R., and Richardson, M. P. (2014). Revealing a Brain Network Endophenotype in Families with Idiopathic Generalised Epilepsy. *PLoS ONE* 9, e110136.

Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-Law Distributions in Empirical Data. *SIAM Review* 51, 661–703.

Coombes, S. (2005). Waves, bumps, and patterns in neural field theories. *Biological Cybernetics* 93, 91–108.

Coombes, S. (2010). Large-scale neural dynamics: Simple and complex. *NeuroImage*. Computational Models of the Brain 52, 731–739.

Cosandier-Rimele, D., Badier, J. M., Chauvel, P., and Wendling, F. (2007). A Physiologically Plausible Spatio-Temporal Model for EEG Signals Recorded With Intracerebral Electrodes in Human Partial Epilepsy. *IEEE Transactions on Biomedical Engineering* 54, 380–388.

Daffertshofer, A. and van Wijk, B. C. (2011). On the influence of amplitude on the connectivity between phases. *Frontiers in neuroinformatics* 5.

Daunizeau, J., David, O., and Stephan, K. E. (2011). Dynamic causal modelling: A critical review of the biophysical and statistical foundations. *NeuroImage* 58, 312–322.

Daunizeau, J., Kiebel, S., and Friston, K. (2009). Dynamic causal modelling of distributed electromagnetic responses. *NeuroImage* 47, 590–601.

David, O. and Friston, K. J. (2003). A neural mass model for MEG/EEG: coupling and neuronal dynamics. *NeuroImage* 20, 1743–1755.

David, O., Harrison, L., and Friston, K. J. (2005). Modelling event-related responses in the brain. *NeuroImage* 25, 756–770.

Deco, G., Jirsa, V. K., Robinson, P. A., Breakspear, M., and Friston, K. (2008). The Dynamic Brain: from spiking neurons to neural masses and cortical fields. *PLoS computational biology* 4, e1000092.

Deco, G., Jirsa, V., McIntosh, A. R., Sporns, O., and Kötter, R. (2009). Key role of coupling, delay, and noise in resting brain fluctuations. *Proceedings of the National Academy of Sciences* 106, 10302–10307.

Delorme, A. and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods* 134, 9–21.

Dhooge, A., Govaerts, W., and Kuznetsov, Y. A. (2003). MATCONT: a MATLAB package for numerical bifurcation analysis of ODEs. *ACM Transactions on Mathematical Software (TOMS)* 29, 141–164.

Dhooge, A., Govaerts, W., Kuznetsov, Y. A., Mestrom, W., Riet, A. M., and Sautois, B. (2006). *MATCONT and CL-MATCONT continuation toolboxes in MATLAB*. Citeseer.

Dumont, G., Northoff, G., and Longtin, A. (2014). Linear noise approximation for oscillations in a stochastic inhibitory network with delay. *Physical Review E* 90, 012702.

Ecker, A. S., Berens, P., Keliris, G. A., Bethge, M., Logothetis, N. K., and Tolias, A. S. (2010). Decorrelated neuronal firing in cortical microcircuits. *Science* 327, 584–587.

Engel, A. K., Fries, P., and Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top–down processing. *Nature Reviews Neuroscience* 2, 704–716.

Ermentrout, B. (1986). Losing amplitude and saving phase. *Nonlinear Oscillations in Biology and Chemistry*. Springer, pp. 98–114.

Ermentrout, G. B. and Kopell, N. (1991). Multiple pulse interactions and averaging in systems of coupled neural oscillators. *Journal of Mathematical Biology* 29, 195–217.

Ermentrout, G. B., Galán, R. F., and Urban, N. N. (2008). Reliability, synchrony and noise. *Trends in Neurosciences* 31, 428–434.

Ermentrout, G. and Kopell, N. (1990). Oscillator Death in Systems of Coupled Neural Oscillators. *SIAM Journal on Applied Mathematics* 50, 125–146.

Faisal, A. A., Selen, L. P. J., and Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience* 9, 292–303.

Faisal, A. A. and Laughlin, S. B. (2007). Stochastic Simulations on the Reliability of Action Potential Propagation in Thin Axons. *PLOS Comput Biol* 3, e79.

Fein, G., Raz, J., Brown, F. F., and Merrin, E. L. (1988). Common reference coherence data are confounded by power and phase effects. *Electroencephalography and Clinical Neurophysiology* 69, 581–584.

Firth, W. J. and Yao, A. M. (2005). Giant excess noise and transient gain in misaligned laser cavities. *Physical review letters* 95, 073903.

Fisher, N. I. (1993). *Statistical Analysis of Circular Data*. Cambridge University Press.

Freeman, W. (1992). Predictions on Neocortical Dynamics Derived from Studies in Paleocortex. *Induced rhythms in the brain*. Ed. by E. Başar and T. H. Bullock. Brain dynamics series. Birkhäuser, pp. 183–199.

Freeman, W. J. (1972). Waves, Pulses, and the Theory of Neural Masses. *Progress in Theoretical Biology* 2.1.

Freeman, W. J. (1975). *Mass Action in the Nervous System*. Citeseer.

Freyer, F., Aquino, K., Robinson, P. A., Ritter, P., and Breakspear, M. (2009). Bistability and non-Gaussian fluctuations in spontaneous cortical activity. *The Journal of Neuroscience* 29, 8512.

Freyer, F., Roberts, J. A., Becker, R., Robinson, P. A., Ritter, P., and Breakspear, M. (2011). Biophysical Mechanisms of Multistability in Resting-State Cortical Rhythms. *The Journal of Neuroscience* 31, 6353.

Freyer, F., Roberts, J. A., Ritter, P., and Breakspear, M. (2012). A Canonical Model of Multistability and Scale-Invariance in Biological Systems. *PLoS Comput Biol* 8, e1002634.

Friston, K. J. and Dolan, R. J. (2010). Computational and dynamic models in neuroimaging. *Neuroimage* 52, 752–765.

Friston, K. J., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. *Neuroimage* 19, 1273–1302.

Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping* 2, 56–78.

Friston, K. J. (2000). The labile brain. II. Transients, complexity and selection. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 355, 237–252.

Friston, K., Breakspear, M., and Deco, G. (2012). Perception and self-organized instability. *Frontiers in Computational Neuroscience* 6, 44.

Gardiner, C. (2010). *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer Berlin Heidelberg.

Gerstner, W. and Kistler, W. M. (2002). *Spiking neuron models: Single neurons, populations, plasticity*. Cambridge university press.

Gollo, L. L. and Breakspear, M. (2014). The frustrated brain: from dynamics on motifs to communities and networks. *Phil. Trans. R. Soc. B* 369, 20130532.

Gollo, L. L., Copelli, M., and Roberts, J. A. (2016). Diversity improves performance in excitable networks. *PeerJ* 4, e1912.

Golubitsky, M., Schaeffer, D. G., and Stewart, I. (1985). *Singularities and groups in bifurcation theory*. New York: Springer-Verlag.

Golubitsky, M., Stewart, I., and Schaeffer, D. G. (1988). *Singularities and Groups in Bifurcation Theory Volume II*. New York, NY: Springer New York.

Goodfellow, M. and Glendinning, P. (2013). Mechanisms of Intermittent State Transitions in a Coupled Heterogeneous Oscillator Model of Epilepsy. *The Journal of Mathematical Neuroscience* 3, 17.

Goodfellow, M., Schindler, K., and Baier, G. (2011). Intermittent spike-wave dynamics in a heterogeneous, spatially extended neural mass model. *NeuroImage* 55, 920–932.

Graham, R. and Haken, H. (1970). Laserlight — first example of a second-order phase transition far away from thermal equilibrium. *Zeitschrift für Physik* 237, 31–46.

Grimbert, F. and Faugeras, O. (2006). *Analysis of Jansen's model of a single cortical column*. Tech. rep. INRIA (Institut National de Recherche en Informatique et en Automatique).

Gröchenig, K. (2001). *Foundations of Time-Frequency Analysis*. Applied and Numerical Harmonic Analysis. Boston, MA: Birkhäuser Boston.

Grossmann, S. (2000). The onset of shear flow turbulence. *Reviews of modern physics* 72, 603.

Guckenheimer, J. and Holmes, P. (1985). *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer Science & Business Media.

Gutkin, B. S., Laing, C. R., Colby, C. L., Chow, C. C., and Ermentrout, G. B. (2001). Turning On and Off with Excitation: The Role of Spike-Timing Asynchrony and Synchrony in Sustained Neural Activity. *Journal of Computational Neuroscience* 11, 121–134.

Haken, H., Kelso, J. A. S., and Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological cybernetics* 51, 347–356.

Haken, H. (1978). *Synergetics: An Introduction Nonequilibrium Phase Transitions and Self-Organization in Physics, Chemistry and Biology*. Vol. 1. Springer Series in Synergetics. Berlin, Heidelberg: Springer Berlin Heidelberg.

Haken, H. (1983). *Advanced Synergetics: Instability Hierarchies of Self-Organizing Systems and Devices*. Vol. 20. Springer Series in Synergetics. Berlin, Heidelberg: Springer Berlin Heidelberg.

Hale, J. K. (1971). *Functional Differential Equations*. Vol. 3. Applied Mathematical Sciences. New York, NY: Springer US.

Hamel, W. and Woerdman, J. (1989). Nonorthogonality of the longitudinal eigenmodes of a laser. *Physical Review A* 40, 2785.

Hansel, D., Mato, G., and Meunier, C. (1993). Phase dynamics for weakly coupled Hodgkin-Huxley neurons. *EPL (Europhysics Letters)* 23, 367.

He, B. J., Zempel, J. M., Snyder, A. Z., and Raichle, M. E. (2010). The temporal structures and functional significance of scale-free brain activity. *Neuron* 66, 353–369.

Heaton, C. and Peake, N. (2007). Transient growth in vortices with axial flow. *Journal of Fluid Mechanics* 587, 271.

Herz, A. V. M., Gollisch, T., Machens, C. K., and Jaeger, D. (2006). Modeling Single-Neuron Dynamics and Computations: A Balance of Detail and Abstraction. *Science* 314, 80–85.

Hodgkin, A. L. and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology* 117, 500–544.

Holgado, A. J. N., Terry, J. R., and Bogacz, R. (2010). Conditions for the generation of beta oscillations in the subthalamic nucleus–globus pallidus network. *The Journal of Neuroscience* 30, 12340–12352.

Honey, C. J., Kötter, R., Breakspear, M., and Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences* 104, 10240–10245.

Hopf, E. (1942). Abzweigung einer periodischen Lösung von einer stationären Lösung eines Differentialsystems. *Ber. Math.-Phys. Kl Sächs. Akad. Wiss. Leipzig* 94, 1–22.

Huang, G., Zhang, D., Meng, J., and Zhu, X. (2011). Interactions between two neural populations: A mechanism of chaos and oscillation in neural mass model. *Neurocomputing* 74, 1026–1034.

Izhikevich, E. M. (2007). *Dynamical systems in neuroscience: the geometry of excitability and bursting*. MIT press.

Jacobs, K. (2010). *Stochastic Processes for Physicists: Understanding Noisy Systems*. Cambridge University Press.

Jansen, B. H. and Rit, V. G. (1995). Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biological Cybernetics* 73, 357–366.

Jansen, B. H., Zouridakis, G., and Brandt, M. E. (1993). A neurophysiologically-based mathematical model of flash visual evoked potentials. *Biological Cybernetics* 68, 275–283.

Jensen, O., Goel, P., Kopell, N., Pohja, M., Hari, R., and Ermentrout, B. (2005). On the human sensorimotor-cortex beta rhythm: sources and modeling. *Neuroimage* 26, 347–355.

Jirsa, V. K. and Haken, H. (1996). Field theory of electromagnetic brain activity. *Physical Review Letters* 77, 960–963.

Jirsa, V. K., Stacey, W. C., Quilichini, P. P., Ivanov, A. I., and Bernard, C. (2014). On the nature of seizure dynamics. *Brain*, awu133.

Jirsa, V. and Müller, V. (2013). Cross-frequency coupling in real and virtual brain networks. *Frontiers in Computational Neuroscience* 7, 78.

Kalitzin, S. N., Velis, D. N., and Lopes da Silva, F. H. (2010). Stimulation-based anticipation and control of state transitions in the epileptic brain. *Epilepsy & Behavior* 17, 310–323.

Kalitzin, S., Koppert, M., Petkov, G., and Da Silva, F. L. (2014). Multiple oscillatory states in models of collective neuronal dynamics. *International Journal of Neural Systems* 24, 1450020.

Kelso, J. A. S. (2010). Instabilities and phase transitions in human brain and behavior. *Frontiers in Human Neuroscience* 4, 1–2.

Kelso, J. A. S., Scholz, J. P., and Schöner, G. (1986). Nonequilibrium phase transitions in coordinated biological motion: Critical fluctuations. *Physics Letters A* 118, 279–284.

Kilner, J. M., Baker, S. N., Salenius, S., Jousmäki, V., Hari, R., and Lemon, R. N. (1999). Task-dependent modulation of 15-30 Hz coherence between rectified EMGs from human hand and forearm muscles. *The Journal of Physiology* 516, 559–570.

Kinfe, T. M. and Vesper, J. (2013). The Impact of Multichannel Microelectrode Recording (MER) in Deep Brain Stimulation of the Basal Ganglia. *Stereotactic and Functional Neurosurgery*. Ed. by G. Nikkhah and M. Pinsker. Acta Neurochirurgica Supplement 117. Springer Vienna, pp. 27–33.

Kinouchi, O. and Copelli, M. (2006). Optimal dynamical range of excitable networks at criticality. *Nature Physics* 2, 348–351.

Kloeden, P. E. and Platen, E. (1992). Higher-order implicit strong numerical schemes for stochastic differential equations. *Journal of Statistical Physics* 66, 283–314.

Kloeden, P. E. and Platen, E. (1999). *Numerical solution of Stochastic Differential Equations*. corrected 3rd printing. Applications of mathematics 23. Berlin: Springer.

Knight, B. W. (2000). Dynamics of encoding in neuron populations: some general mathematical features. *Neural Computation* 12, 473–518.

Knight, B. W., Manin, D., and Sirovich, L. (1996). Dynamical models of interacting neuron populations in visual cortex. Vol. 54, pp. 4–8.

Kopell, N., Ermentrout, G. B., Whittington, M. A., and Traub, R. D. (2000). Gamma Rhythms and Beta Rhythms Have Different Synchronization Properties. *Proceedings of the National Academy of Sciences of the United States of America* 97, 1867–1872.

Kopell, N., Kramer, M., Malerba, P., and Whittington, M. (2010). Are different rhythms good for different functions? *Frontiers in human neuroscience* 4, 1–9.

Koppert, M. M. J., Kalitzin, S., Silva, F. L. d., and Viergever, M. A. (2012). Connectivity and phase coherence in neural network models of interconnected Z4-bi-stable units. *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5458–5461.

Koppert, M., Kalitzin, S., Velis, D., Lopes Da Silva, F., and Viergever, M. A. (2014). Dynamics of collective multi-stability in models of multi-unit neuronal systems. *International Journal of Neural Systems* 24, 1430004.

Kuehn, C. (2011). A mathematical framework for critical transitions: Bifurcations, fast-slow systems and stochastic dynamics. *Physica D: Nonlinear Phenomena* 240, 1020–1035.

Kunze, T., Hunold, A., Haueisen, J., Jirsa, V., and Spiegler, A. (2016). Transcranial direct current stimulation changes resting state functional connectivity: A large-scale brain network modeling study. *NeuroImage* 140, 174–187.

Kuramoto, Y. (1984). Cooperative Dynamics of Oscillator Community A Study Based on Lattice of Rings. *Progress of Theoretical Physics Supplement* 79, 223–240.

Kuramoto, Y. and Haken, H. (1984). *Chemical Oscillations, Waves, and Turbulence*. Vol. 19. Springer Series in Synergetics. Berlin, Heidelberg: Springer Berlin Heidelberg.

Kuznetsov, Y. A. (2010). *Elements of Applied Bifurcation Theory*. 3rd. Applied mathematical sciences 112. New York, NY: Springer.

Lachaux, J., Rodriguez, E., Martinerie, J., Varela, F. J., et al. (1999). Measuring phase synchrony in brain signals. *Human brain mapping* 8, 194–208.

Laing, C. R., Frewen, T., and Kevrekidis, I. G. (2010). Reduced models for binocular rivalry. *Journal of Computational Neuroscience* 28, 459–476.

Linkenkaer-Hansen, K., Nikouline, V. V., Palva, J. M., and Ilmoniemi, R. J. (2001). Long-range temporal correlations and scaling behavior in human brain oscillations. *The Journal of neuroscience* 21, 1370.

Linkenkaer-Hansen, K., Nikulin, V. V., Palva, J. M., Kaila, K., and Ilmoniemi, R. J. (2004). Stimulus-induced change in long-range temporal correlations and scaling behaviour of sensorimotor oscillations. *European Journal of Neuroscience* 19, 203–218.

Lopes da Silva, F. H., Hoeks, A., Smits, H., and Zetterberg, L. H. (1974). Model of brain rhythmic activity. *Biological Cybernetics* 15, 27–37.

Lopes da Silva, F. H., Pijn, J. P., Velis, D., and Nijssen, P. C. G. (1997). Alpha rhythms: noise, dynamics and models. *International Journal of Psychophysiology* 26, 237–249.

Lopes da Silva, F. and Storm Van Leeuwen, W. (1977). The cortical source of the alpha rhythm. *Neuroscience Letters* 6, 237–241.

Markram, H. (2006). The Blue Brain Project. *Nature Reviews Neuroscience* 7, 153–160.

Marreiros, A. C., Daunizeau, J., Kiebel, S. J., and Friston, K. J. (2008). Population dynamics: variance and the sigmoid activation function. *Neuroimage* 42, 147–157.

Marsden, J. E. and McCracken, M. (1976). *The Hopf Bifurcation and Its Applications*. Vol. 19. Applied Mathematical Sciences. New York, NY: Springer New York.

Marten, F., Rodrigues, S., Benjamin, O., Richardson, M. P., and Terry, J. R. (2009). Onset of polyspike complexes in a mean-field model of human electroencephalography and its application to absence epilepsy. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 367, 1145–1161.

Maruyama, G. (1955). Continuous Markov processes and stochastic equations. *Rendiconti del Circolo Matematico di Palermo* 4, 48.

Mehrkanoon, S., Breakspear, M., and Boonstra, T. W. (2014). The reorganization of corticomuscular coherence during a transition between sensorimotor states. *NeuroImage* 100, 692–702.

Mehrkanoon, S., Mehrkanoon, S., Breakspear, M., Daffertshofer, A., and Boonstra, T. (2011). Generalized time-frequency coherency for assessing neural interactions in electrophysiological recordings. *Nature Precedings*.

Miller, K. J., Sorensen, L. B., Ojemann, J. G., and Nijs, M. d. (2009). Power-law scaling in the brain surface electric potential. *PLoS computational biology* 5, e1000609.

Molaee-Ardekani, B., Márquez-Ruiz, J., Merlet, I., Leal-Campanario, R., Gruart, A., Sánchez-Campusano, R., Birot, G., Ruffini, G., Delgado-García, J., and Wendling, F. (2013). Effects of transcranial Direct Current Stimulation (tDCS) on cortical activity: A computational modeling study. *Brain Stimulation* 6, 25–39.

Moran, R. J., Kiebel, S. J., Stephan, K. E., Reilly, R. B., Daunizeau, J., and Friston, K. J. (2007). A neural mass model of spectral responses in electrophysiology. *NeuroImage* 37, 706–720.

Mormann, F., Kreuz, T., Rieke, C., Andrzejak, R. G., Kraskov, A., David, P., Elger, C. E., and Lehnertz, K. (2005). On the predictability of epileptic seizures. *Clinical Neurophysiology* 116, 569–587.

Murdock, J. A. (1999). *Perturbations: Theory and Methods*. 2nd ed. SIAM.

Murdock, J. A. (2003). *Normal Forms and Unfoldings for Local Dynamical Systems*. Springer Science & Business Media.

New, G. (1995). The origin of excess noise. *Journal of Modern Optics* 42, 799–810.

Niedermeyer, E. and Silva, F. H. L. d. (2005). *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams & Wilkins.

Nunez, P. L. (1974). The brain wave equation: a model for the EEG. *Mathematical Biosciences* 21, 279–297.

Nunez, P. L. (1995). *Neocortical Dynamics and Human EEG Rhythms*. Oxford University Press.

Nunez, P. L. and Srinivasan, R. (2006). *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press.

Nunez, P. L., Wingeier, B. M., and Silberstein, R. B. (2001). Spatial-temporal structures of human alpha rhythms: Theory, microcurrent sources, multiscale measurements, and global binding of local networks. *Human Brain Mapping* 13, 125–164.

Nuwer, M. R., Comi, G., Emerson, R., Fuglsang-Frederiksen, A., Guérit, J., Hinrichs, H., Ikeda, A., Jose C. Luccas, F., and Rappelsburger, P. (1998). IFCN standards for digital recording of clinical EEG. *Electroencephalography and Clinical Neurophysiology* 106, 259–261.

Omurtag, A., Knight, B., and Sirovich, L. (2000). On the Simulation of Large Populations of Neurons. *Journal of Computational Neuroscience* 8, 51–63.

Orfanidis, S. J. (1996). *Optimum Signal Processing*. 2nd. New York: McGraw-Hill.

Parlett, D. B. N. and Reinsch, D. C. (1969). Balancing a matrix for calculation of eigenvalues and eigenvectors. *Numerische Mathematik* 13, 293–304.

Pascual-Marqui, R. D. and Lehamann, D. (1993). Topographic maps, source localization inference, and the reference electrode: comments on a paper by Desmedt et al. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section* 88, 532–533.

Pascual-Marqui, R. D., Michel, C. M., and Lehmann, D. (1994). Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *International Journal of Psychophysiology* 18, 49–65.

Petkov, G., Goodfellow, M., Richardson, M. P., and Terry, J. R. (2014). A critical role for network structure in seizure onset: a computational modeling approach. *Epilepsy* 5, 261.

Pikovsky, A., Rosenblum, M., and Kurths, J. (2003). *Synchronization: A Universal Concept in Nonlinear Sciences*. Cambridge Nonlinear Science Series. Cambridge University Press.

Poincaré, H. (1886). On the irregular integrals of linear equations. *Acta Math* 8, 295.

Poincaré, H. (1879). *Sur les propriétés des fonctions définies par les équations aux différences partielles*. 4. Gauthier-Villars.

Poincaré, H. (1892). *Les méthodes nouvelles de la mécanique céleste*. Vol. 1. Gauthier-Villars.

Pomeau, Y. and Manneville, P. (1980). Intermittent transition to turbulence in dissipative dynamical systems. *Communications in Mathematical Physics* 74, 189–197.

Popovych, O. V. and Tass, P. A. (2011). Macroscopic entrainment of periodically forced oscillatory ensembles. *Progress in Biophysics and Molecular Biology*. BrainModes: The role of neuronal oscillations in health and disease 105, 98–108.

Priesemann, V., Valderrama, M., Wibral, M., and Quyen, M. L. V. (2013). Neuronal Avalanches Differ from Wakefulness to Deep Sleep – Evidence from Intracranial Depth Recordings in Humans. *PLOS Comput Biol* 9, e1002985.

Risken, H. (1996). *The Fokker-Planck Equation: Methods of Solution and Applications*. 2nd ed. Springer series in synergetics v. 18. New York: Springer-Verlag.

Roberts, A. J. (2007). *Computer algebra derives normal forms of stochastic differential equations*. Tech. rep. University of Southern Queensland, Faculty of Sciences, Department of Maths and Computing.

Roberts, A. J. (2008). Normal form transforms separate slow and fast modes in stochastic dynamical systems. *Physica A: Statistical Mechanics and Its Applications* 387, 12–38.

Roberts, A. J. (2009). *Normal form of stochastic or deterministic multiscale differential equations*. Tech. rep.

Roberts, J. A. and Robinson, P. A. (2012). Corticothalamic dynamics: Structure of parameter space, spectra, instabilities, and reduced model. *Physical Review E* 85, 011910.

Robinson, P. A., Rennie, C. J., Phillips, A. J. K., Kim, J. W., and Roberts, J. A. (2010). Phase transitions in physiologically-based multiscale mean-field brain models. *Modeling Phase Transitions in the Brain*. Springer, pp. 179–201.

Robinson, P. A., Rennie, C. J., and Rowe, D. L. (2002). Dynamics of large-scale brain activity in normal arousal states and epileptic seizures. *Physical Review E* 65, 041924.

Robinson, P. A., Rennie, C. J., and Wright, J. J. (1997). Propagation and stability of waves of electrical activity in the cerebral cortex. *Physical Review E* 56, 826.

Robinson, P., Rennie, C., Rowe, D., and O'Connor, S. (2004). Estimation of multiscale neurophysiologic parameters by electroencephalographic means. *Human brain mapping* 23, 53–72.

Robinson, P., Rennie, C., Wright, J., Bahramali, H., Gordon, E., and Rowe, D. (2001). Prediction of electroencephalographic spectra from neurophysiology. *Physical Review E* 63, 021903.

Rodrigues, S., Chizhov, A. V., Marten, F., and Terry, J. R. (2010). Mappings between a macroscopic neural-mass model and a reduced conductance-based model. *Biological cybernetics* 102, 361–371.

Rößler, A. (2010). Runge-Kutta Methods for the Strong Approximation of Solutions of Stochastic Differential Equations. *SIAM Journal on Numerical Analysis* 48, 922–952.

Rowat, P. F. and Greenwood, P. E. (2011). Identification and Continuity of the Distributions of Burst-Length and Interspike Intervals in the Stochastic Morris-Lecar Neuron. *Neural Computation* 23, 1–31.

Rubehn, B., Fries, P., and Stieglitz, T. (2009). MEMS-Technology for Large-Scale, Multichannel ECoG-Electrode Array Manufacturing. *4th European Conference of the International Federation for Medical and Biological Engineering*. Ed. by J. V. Sloten, P. Verdonck, M. Nyssen, and J. Haueisen. IFMBE Proceedings 22. Springer Berlin Heidelberg, pp. 2413–2416.

Rubinov, M., Sporns, O., Thivierge, J., and Breakspear, M. (2011). Neurobiologically Realistic Determinants of Self-Organized Criticality in Networks of Spiking Neurons. *PLOS Comput Biol* 7, e1002038.

Rümelin, W. (1982). Numerical treatment of stochastic differential equations. *SIAM Journal on Numerical Analysis* 19, 604–613.

Sanz-Leon, P., Knock, S. A., Spiegler, A., and Jirsa, V. K. (2015). Mathematical framework for large-scale brain network modeling in The Virtual Brain. *NeuroImage* 111, 385–430.

Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., Held, H., Nes, E. H. v., Rietkerk, M., and Sugihara, G. (2009). Early-warning signals for critical transitions. *Nature* 461, 53–59.

Schmidt, H., Petkov, G., Richardson, M. P., and Terry, J. R. (2014). Dynamics on Networks: The Role of Local Dynamics and Global Networks on the Emergence of Hypersynchronous Neural Activity. *PLoS Comput Biol* 10, e1003947.

Schmidt, H., Woldman, W., Goodfellow, M., Chowdhury, F. A., Koutroumanidis, M., Jewell, S., Richardson, M. P., and Terry, J. R. (2016). A computational biomarker of idiopathic generalized epilepsy from resting state EEG. *Epilepsia* 57, e200–e204.

Scholz, J. P., Kelso, J. A. S., and Schöner, G. (1987). Nonequilibrium phase transitions in coordinated biological motion: critical slowing down and switching time. *Physics Letters A* 123, 390–394.

Schuster, H. G. and Wagner, P. (1990). A model for neuronal oscillations in the visual cortex. *Biological Cybernetics* 64, 77–82.

Seth, A. K., Barrett, A. B., and Barnett, L. (2015). Granger Causality Analysis in Neuroscience and Neuroimaging. *Journal of Neuroscience* 35, 3293–3297.

Shackman, A. J., McMenamin, B. W., Maxwell, J. S., Greischar, L. L., and Davidson, R. J. (2010). Identifying robust and sensitive frequency bands for interrogating neural oscillations. *NeuroImage* 51, 1319–1333.

Shew, W. L. and Plenz, D. (2013). The Functional Benefits of Criticality in the Cortex. *The Neuroscientist* 19, 88–100.

Shriki, O., Hansel, D., and Sompolinsky, H. (2003). Rate models for conductance-based cortical neuronal networks. *Neural computation* 15, 1809–1841.

Siegman, A. (1989). Excess spontaneous emission in non-Hermitian optical systems. I. Laser amplifiers. *Physical Review A* 39, 1253.

Sokolov, I. M. and Klafter, J. (2005). From diffusion to anomalous diffusion: A century after Einstein's Brownian motion. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 15, 026103–026103.

Sotero, R. C., Trujillo-Barreto, N. J., Iturria-Medina, Y., Carbonell, F., and Jimenez, J. C. (2007). Realistically coupled neural mass models can generate EEG rhythms. *Neural computation* 19, 478–512.

Spiegler, A. and Jirsa, V. (2013). Systematic approximations of neural fields through networks of neural masses in the virtual brain. *NeuroImage* 83, 704–725.

Spiegler, A., Kiebel, S. J., Atay, F. M., and Knösche, T. R. (2010). Bifurcation analysis of neural mass models: Impact of extrinsic inputs and dendritic time constants. *NeuroImage* 52, 1041–1058.

Stam, C. J. and de Bruin, E. A. (2004). Scale-free dynamics of global functional connectivity in the human brain. *Human brain mapping* 22, 97–109.

Stam, C. J., Pijn, J. P. M., Suffczynski, P., and Silva, F. H. L. d. (1999). Dynamics of the human alpha rhythm: evidence for non-linearity? *Clinical Neurophysiology* 110, 1801–1813.

Stam, C. J., Nolte, G., and Daffertshofer, A. (2007). Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. *Human Brain Mapping* 28, 1178–1193.

Stefanescu, R. A. and Jirsa, V. K. (2008). A low dimensional description of globally coupled heterogeneous neural networks of excitatory and inhibitory neurons. *PLoS Computational Biology* 4, e1000219.

Steyn-Ross, D. A., Steyn-Ross, M. L., Wilson, M. T., and Sleigh, J. W. (2006). White-noise susceptibility and critical slowing in neurons near spiking threshold. *Physical Review E* 74, 051920.

Strogatz, S. H. (2000). From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D: Nonlinear Phenomena* 143, 1–20.

Tass, P., Rosenblum, M. G., Weule, J., Kurths, J., Pikovsky, A., Volkmann, J., Schnitzler, A., and Freund, H. (1998). Detection of n: m phase locking from noisy data: application to magnetoencephalography. *Physical review letters* 81, 3291.

Tchumatchenko, T., Geisel, T., Volgushev, M., and Wolf, F. (2011). Spike correlations–what can they tell about synchrony? *Frontiers in neuroscience* 5.

Terman, D., Rubin, J. E., Yew, A. C., and Wilson, C. J. (2002). Activity Patterns in a Model for the Subthalamopallidal Network of the Basal Ganglia. *Journal of Neuroscience* 22, 2963–2976.

Terry, J., Benjamin, O., and Richardson, M. (2012). Seizure generation: The role of nodes and networks. *Epilepsia* 53.

Torre, V. (1974). Synchronization of non-linear biochemical oscillators coupled by diffusion. *Biological Cybernetics* 17, 137–144.

Touboul, J. and Destexhe, A. (2010). Can Power-Law Scaling and Neuronal Avalanches Arise from Stochastic Dynamics? *PLOS ONE* 5, e8982.

Troy, W. C. and McLeod, J. B. (1976). Oscillation Phenomena in the Hodgkin-Huxley Equations. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics* 74, 299–310.

Valdes-Sosa, P. A., Roebroeck, A., Daunizeau, J., and Friston, K. (2011). Effective connectivity: Influence, causality and biophysical modeling. *NeuroImage* 58, 339–361.

van Albada, S. J., Kerr, C. C., Chiang, A. K. I., Rennie, C. J., and Robinson, P. A. (2010). Neurophysiological changes with age probed by inverse modeling of EEG spectra. *Clinical Neurophysiology* 121, 21–38.

van Albada, S. J. and Robinson, P. A. (2009). Mean-field modeling of the basal ganglia-thalamocortical system. I: Firing rates in healthy and parkinsonian states. *Journal of Theoretical Biology* 257, 642–663.

van Kampen, N. G. (2011). *Stochastic Processes in Physics and Chemistry*. 3rd ed. Burlington: Elsevier Science.

Vinck, M., Oostenveld, R., van Wingerden, M., Battaglia, F., and Pennartz, C. M. A. (2011). An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias. *NeuroImage* 55, 1548–1565.

Vindiola, M. M., Vettel, J. M., Gordon, S. M., Franaszczuk, P. J., and McDowell, K. (2014). Applying EEG phase synchronization measures to non-linearly coupled neural mass models. *Journal of Neuroscience Methods* 226, 1–14.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society* 57, 307–333.

Welch, P. (1967). The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics* 15, 70–73.

Wendling, F., Bartolomei, F., Bellanger, J. J., and Chauvel, P. (2002). Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition. *European Journal of Neuroscience* 15, 1499–1508.

Wendling, F., Bellanger, J. J., Bartolomei, F., and Chauvel, P. (2000). Relevance of nonlinear lumped-parameter models in the analysis of depth-EEG epileptic signals. *Biological Cybernetics* 83, 367–378.

Wendling, F., Benquet, P., Bartolomei, F., and Jirsa, V. (2016). Computational models of epileptiform activity. *Journal of Neuroscience Methods*. Methods and Models in Epilepsy Research 260, 233–251.

Wiktorsson, M. (2001). Joint Characteristic Function and Simultaneous Simulation of Iterated Itô Integrals for Multiple Independent Brownian Motions. *The Annals of Applied Probability* 11, 470–487.

Wilson, H. R. and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical journal* 12, 1–24.

Winfree, A. T. (1967). Biological rhythms and the behavior of populations of coupled oscillators. *Journal of Theoretical Biology* 16, 15–42.

Womelsdorf, T., Schoffelen, J., Oostenveld, R., Singer, W., Desimone, R., Engel, A. K., and Fries, P. (2007). Modulation of Neuronal Interactions Through Neuronal Synchronization. *Science* 316, 1609–1612.

Wong, K. and Wang, X. (2006). A recurrent network mechanism of time integration in perceptual decisions. *The Journal of neuroscience* 26, 1314–1328.

Wright, J. J. and Liley, D. T. J. (1994). A millimetric-scale simulation of electrocortical wave dynamics based on anatomical estimates of cortical synaptic density. *Network: Computation in Neural Systems* 5, 191–202.

Yang, Y., Solis-Escalante, T., Yao, J., Daffertshofer, A., Schouten, A. C., and van der Helm, F. C. T. (2015). A General Approach for Quantifying Nonlinear Connectivity in the Nervous System Based on Phase Coupling. *International Journal of Neural Systems* 26, 1550031.

Zandt, B., Visser, S., van Putten, M. J. A. M., and ten Haken, B. (2014). A neural mass model based on single cell dynamics to model pathophysiology. *Journal of Computational Neuroscience* 37, 549–568.

Zavaglia, M., Astolfi, L., Babiloni, F., and Ursino, M. (2006). A neural mass model for the simulation of cortical activity estimated from high resolution EEG during cognitive or motor tasks. *Journal of neuroscience methods* 157, 317–329.